1991

A potential analysis of internal migration in the U.S.

Man-keung Lee
Iowa State University

A Potential Analysis of Internal Migration in the U.S.

by

Man-keung Lee

A Thesis Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

Department: Economics
Major: Economics

Iowa State University
Ames, Iowa
1991

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Urban economists, geographers, demographers, and many other researchers recognized that a big shift of U.S. population occurred in 1970s. A paper written by Greenwood and Hunt [8] clearly stated the direction of shift was from North east and North central to the South and West. This shift created new cities and new regional economics. In addition, it also created new shifts of Federal and state governments' resources. Thus, the government needs to study the shift of migration flow. The government needs to answer why the people moved from one place to another. What are the motives for the move? How many of them will move ... ? Some of these questions will be answered in this study.

In this study, three economic variables, regional population (P), per capita real income (I) and unemployment rate (U) are chosen to explain the economic aspects of the variations of the per capita in-migration flow (M/P), where M is the internal in-migration flows. The use of total measures in this study has been beliefly described in the end of this paper, see section "Notes".

Gravity model had been frequently used for spatial interaction analysis for many years. Potential analysis, another application of theory from physics to socialeconomics study, was used to describe the forces applied on spatial interactions [25], [24]. Potentials are restricted to describe facts. This study is planned to introduce the application of potential analysis on U.S. internal in-migration flows in an econometric model.

The third purpose is to try to use certain information to set up the econometric model. Then, it needs some certain information to predict the future in-migration flows. Therefore, this model can overcome the needs of predicted inputs in many other prediction models.

2. PREVIOUS EMPIRICAL STUDIES

In this chapter, three studies on related topics will be discussed. The first paper to be discussed is done by M. J. Greenwood [7], while the second one is done by J. R. Kau and S. F. Sirmans [12] and the third one by P. Mueser [14].

2.1 Greenwood's Study

Greenwood set up a 9-equation simultaneous equation system with another five identities to regress 14 endogenous variables. There were almost eleven variables in each equation including four regional dummy variables. His study covered the time span from 1950 to 1970 and focused on the Civilian Labor Force (CLF) in-migration of the Standard Metropolitan Statistical Areas (SMSAs). In his study,

$$\begin{aligned} IMFS &= f_3(O\widehat{MT}M, \Delta I\widehat{NC}, \Delta E\widehat{M}P, \Delta N\widehat{EM}P, INC, \\ &\quad UNR, CLF, D1, D2, D3, D4, e3). \\ IMFN &= f_4(O\widehat{MT}M, ..., e4). \end{aligned}$$

where, IM = rate of CLF in In-migration;

IMFS = IM from other SMSAs;

IMFN = IM from nonmetropolitan areas;

$O\widehat{MT}N$ = estimate of out-migration to nonmetropolitan area;

$\Delta\widehat{INC}$ = estimate of rate of income growth;

$\Delta\widehat{EMP}$ = estimate of rate of employment growth;

$\Delta\widehat{UNEMP}$ = estimate of rate of unemployment growth;

INC = income;

UNR = unemployment rates;

CLF = civil labor force; and

D_i = regional dummy variables; $i = 1, ..., 4$.

The rates of growth mentioned for 1955-1960 and 1965-1970 CLF movements are referred to the changes that occurred between 1950-1960 and 1960-1970 publications of Census migration data.

He sampled 63 common SMSAs during 1960 and 1970. He could explain about 94 (93) percent of the variation of the CLF in-migration behavior of those 63 SMSAs from other SMSAs during 1950-1960 (1960-1970). But he could explain about 91 (75) percent of that from Nonmetropolitan areas during 1950-1960 (1960-1970). He also pooled the data together and obtained 92 and 85 percent of in-migration effects from different sources.

He did well to focus on the SMSAs with common characteristics instead of on the state base analysis. However, the problem is whether the sample of 63 SMSA's is representative of over 200 SMSAs. Since the samples are not randomly picked, his results are questionable. It is analogous to picking Iowa's agricultural performance to analyse the agricultural performance of the whole U.S. Is the sample large enough to represent the whole population? Another problem is the simultaneous equation system made the regression of in-migration cannot stand up alone. It becomes unclear for the effect of a particular regressor to influence the dependent variable. For instance, the coefficients of the ΔUNEMP and UNR60 (the UNR in 1960) were all positive, 22.27 and 0.178 respectively in 1960 to 1970, do they imply that an increase in the unemployment or unemployment rate in 1960, the in-migration from nonmetropolitan areas could increase, i.e., unemployment induced in-migration? It is hard to accept. He introduced a new suggestion that in-migration may be due to not just relative high income but the high growth of it.

2.2 Kau and Sirmans' Study

Kau and Sirmans' paper was published in 1979. Their paper hypothesized that people try to maximize utility by moving. They made use of the states' data and tried to compare the regression performance by traditional model and recursive model[1]. They ran regressions with 1940 to 1970 data on eight regressors (including lagged in-migration stocks, 2 Stage Least Square Estimates of median family income of origins and destinations, absolute deviation of mean yearly temperature from 65 in origins and destinations, median education level of origins, median age of origins' population, and highway mileage between the highest populated city of origin and destination). They concluded that the recursive model was much better than that by using the traditional model. For instance, the adjusted R^2 (or $\bar{R}^2$) was only 0.628 in 1970 regression by using traditional model, but it increased to 0.916 by using the recursive model.

If we only look at the increment of the $\bar{R}^2$, there was a significant improvement

[1] Recursive system was developed by Wold. It is a special case of the full structural model led by a set of restrictions. The recursive system is the restricted and ordered structural system so that each equation in the system has only one endogenous variable. Then, the whole system can be solved step by step.

by using the recursive model, but after reviewing his regression model carefully, some questionable facts are discovered. First of all, the idea of introducing past migration for the information hypothesis to induce present migration was a good idea, but in the two models he compared, the Migration stocks (MS) was used in the traditional model and the Migration rates (M) was used in the recursive model. Migration stocks are the past migrant who were born in one state and enumerated to another state within a specific period and the migration rates are the number of migrants who were residing in state a in year t and had migrated to state b in year t' then divided by the total population at origin a of year t. Did they assume the MS and M are the same? But MS and M should not have the same explanatory ability, because they were two different variables by definition. It is no good to compare two different methods by using different variables. Secondly, he used 8 variables in the traditional model, and added 3 more lagged dependent variables in the recursive model. Of course, the $\bar{R}^2$ was not necessarily increased but it did in their study. The question was the involvement of insignificant variables. For all the four regressions of the traditional model, almost all coefficients of variables except that of the variable "age" were significantly different from zero at 0.05 level, but some coefficients became insignificant after adding the lagged dependent variables in the recursive model. For example, comparing the 1970 regressions, the coefficients were all statistically significantly different from zero by using traditional model, but the coefficients of "income destination", "age", "temperature destination" and "Migration Stock of 1930" became insignificantly different from zero at even 0.1 level. Although statistical significance sometimes cannot reflect economic significance (We have to consider the economic reasonings of the variables.), but the significance of a variable from zero in a model can provide us a basic idea for the probability of its explanatory ability in repeated sampling. That means those insigificant variables might not affect the migration flow in the 1970. Was that true the migration or migration rate depend mainly on the past migration performances of ten, twenty, thirty, or forty years ago? Didn't the migrants consider the income they could earn after moving? Their study might be over-emphasized the past migration, and so neglected the considerable effects of some other significant variables, such as income and population of destinations. On the other hand, the involvement of too many past migration terms might bring up multi-collinearity problems among the explanatory variables in the recursive model. Some questions were pointed out in their paper but had not been concerned in their study.

2.3 Mueser's Study

Mueser compared seven models using different distance measurements. His second model used geographic distances, r_{ij}, by measuring distance in miles the reported latitude and longitude coordinates of states' largest cities. He also made an assumption of logarithmic transformation of all variables to fit ordinary least square estimation assumptions. By choosing the best fitted variables, the "sending propensity" of location i, G_i and "draw" of location j, K_j, to explain the migration stream between location i and j, M_{ij}, he obtained the following general spatial interaction model:

$$M_{ij} = G_i \cdot K_j \cdot d_{ij}$$

where, d_{ij} indicates the relationship between locations i and j. It can be called the distance or separation effect, so d_{ij} can be seen as a function of r_{ij} (or r_{ij}^{B}). We can expect B to be negative.

For his second model,

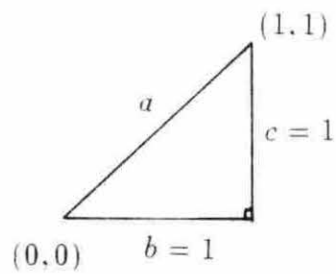
$$\ln M_{ij} = \ln G_i + \ln K_j + B \cdot \ln r_{ij} + \ln e_{ij}.$$

By inputing the states' data, his findings were: (i) the distance elasticity (B) equal -1.2 to -1.3; (ii) R^2 were 0.9019, 0.9097 and 0.9122 for regressions of ln M_{ij} 1955-60, 1965-70 and 1975-80 respectively; and (iii) comparing his work and Greenwood's work in 1969, Greenwood used gravity model with Ordinary Least Squares to estimate distance elasticity, downward bias estimates might be obtained by using geographic distances in gravity model. Two things can be improved from his study to make it more comparable to the behavior in the real world. First, transportation distances (as used by Kau and Sirmans) might be used to substitute for the geographic distance measures. It is because the geographic distances were the shortest possible distances between two places and were measured neglecting the geographic conditions on the earth surface, such as mountains and rivers. The coordinates' distances were usually somewhat under-estimated the actual transportation distances from one place to another unless the two places are located on a very straight highway. If the highways are not built straightly linked every location, then, the following three example cases might show some deviation of the two measures.

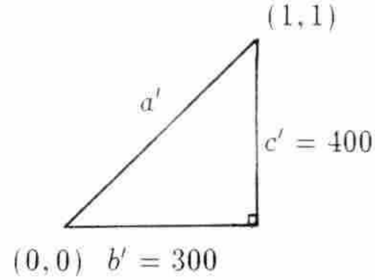
Case 1. For small scale:

geographic distance=a=1.4

transportation distance=b+c=2



geographic distance error=-0.6

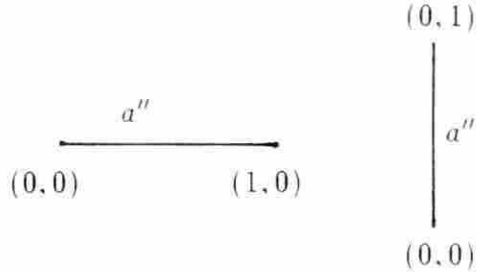


Case 2. For larger scale:

geographic distance=$a' = 500$

transportation distance=$b' + c' = 700$

geographic distance error=-200



Case 3. For minimum error:

geographic distance=a''

transportation distance=a''

geographic distance error=0

Secondly, the structural form of the model need not to be fixed at logarithmic transformation. It is because the best fitted variables chosen according to the logarithmic structural model may be no longer the best fitted variables for the other structural forms. He forced the nonlinear transformed variables to fit a linear model. So, the best structural form and the best chosen variables are interdependent to form best fitted regression line to the dependent variable. The best chosen variables by using his method is only the best for the logarithmic transformation. Therefore, choosing variables first and then selecting the best fitted model structure may be an alternative method to be used.

2.4 Methodology of This Study

This study can be roughly divided into eight procedures. Those procedures are briefly described as the followings:

(1) Define Control Regions and obtain data for the year 1960, 1970 and 1980[2].

(2) Potential transformation of explanatory variables.

(3) Obtain per capita In-migration.

(4) Regress the pooled data (60, 70, 80) to choose the best basic functional form, and verify by Box-Cox Regression.

(5) Transform variables (according to the result in (4)).

(6) Apply different approaches of initial model structures, choose the best initial structure.

(7) Refine the best initial structure (according to the result in (6)), choose the best model.

(8) Make predictions for the 1985 to 1990 In-migration flows.

Detailed descriptions of each procedure and their results are presented in the following chapters. Procedures 1 to 3 are presented in Chapter 4; Procedures 4 to 6 are presented in Chapter 5. Procedures 7 and 8 are presented in Chapter 6. A brief review of potential analysis is presented in Chapter 3. Chapters 7 and 8 are the limitations and conclusion of this study, respectively.

---

[2] The detailed definations of the data will be discussed in Chapter 4.

3. THEORETICAL BACKGROUND

For the urban economists, human geographers or sociologists, gravity and potential models are generally accepted to be used in the studies of spatial interactions. Potential is based on regional interactions and distance restrictions among the regions. It forms a system of regions. The first part of this chapter provides brief notes for the development of potentials. The second part of this chapter reviews some of the applications of potentials.

3.1 The Development of Potentials

According to Carrothers [3] and Isard [9], the earliest gravity concept of human interaction was used by H. C. Carey who applied the concept of molecular gravitation in physics to social science. Human interactions, such as trip volume and population distribution, can be modeled from the socialphysics point of view as,

$$I_{ij} = \frac{\mathrm{f}(\mathrm{P_i},\mathrm{P_j})}{\mathrm{h}(\mathrm{D_{ij}})} \tag{3.1}$$

where, P_i, P_j = size or mass of areas i and j, respectively;

I_{ij} = interaction between center i and center j; and

D_{ij} = distance between center i and center j.

The interaction is hypothesized to be a relationship between a function of the masses in the regions i and j, and a function of the frictional term, distance, between the two regions.

The function $h(D_{ij})$ was used as the square of D_{ij} by E. C. Young to measure migration in the late 1920s. Then, a special case, the force of interaction, was developed by J. Q. Stewart and G. K. Zipf, can be measured as,

$$F_{ij} = G \cdot \frac{P_i \cdot P_j}{D_{ij}^2}$$

where F_{ij}= force of interaction between center i and center j;

G = gravitational constant; and

P_i, P_j = masses of center i and center j, respectively.

Stewart modified the concept to form the gravitational potential produced at region i by a mass at region j, ${}_iV_j$ as,

$${}_iV_j = G \cdot \frac{P_j}{D_{ij}}.$$

The total potential of a region i, ${}_iV$, can be measured as the sum of all gravitational potentials produced at the regions which interact with i, i.e.,

$${}_iV = G \cdot \sum_{j=1}^{n} \frac{P_j}{D_{ij}}$$

where n = number of regions under consideration; and

$j = 1, 2, \ldots, n$.

Other forms of gravitational models are developed afterward by adding weighting factors to the masses or by adding exponential factors to the masses and to the distance. (Readers may refer to Carrothers [3] and Isard [9]).

3.2 Application of Potentials

The earliest development of potentials were focused on the studies of population distributions and demographic patterns (see Stewart [19]). Later on, broader uses of potentials in many forms appeared in many studies. For instance, income potentials (has unit $1,000,000,000. per 100 miles, on average for the period 1940 to 1949) which used total income as mass were calculated by Warntz [24], [25]. In his paper, wheat supply, potato supply, onion supply and strawberry supply space potentials were also produced so called "Product Supply Space Potentials" to help analyse the supply of crops per unit of area. Another "Product Supply Time Potentials" can express the supply of crops at every point of time.

In a discussion of Warntz's paper by Voorhees [23], he stated that Warntz's technique might be possible to solve the problems in urban transportation of trying to determine future travel patterns. Of course, by knowing where people will shop and commute can help better city planning and improve the efficiency of human activities. Further comment by Jurkat [11] on Warntz's potentials suggested that a combined analysis of Product Supply Space Potentials and Income Potential can help to explain the price structures of those products.

Anderson [2] tried to use potential analysis to explain the variations in population density among small areas within a big region. He used the relative density in terms of a function of some parameters to describe the state of the density at any given time. By giving the parameters at different time span, the change in the density of a given area can be estimated. This can be used for substituting the interactions and then measured the total potential of relative density of the small areas. He concluded that " large cities tend to be located at about the same distance from the next larger city ".

Dunn [5] discussed how to apply the potential concept to market analysis. Following Harris's market potential analysis, Dunn formed the market potentials and the transport costs contour maps for Florida in 1948. From the combined analyses of market potential contours and transport cost contours, he suggested to use an index and to compare costs and benefits so that the optimal location of business could be found.

Carrothers [4] made an intensive study using income potentials to predict population trends. He used personal income as the masses of the regions and used rail distances as the distance measures among regions. He calculated three relative income potentials by substituting the weight fall-off function of distance measure with: (i) the rail distance; (ii) the squared rail distance and (iii) the square root of the rail distance[1]. He divided the 48 states into 31 regions. By comparing the arithmetic mean of the forecast error, he found that the distance function of form (ii) was the best (or least error) for the test period 1920-30, form (i) best for 1930-40 and form (iii) for 1940-50.

In this study, potentials (Population Potentials (Pp), Income Potentials (Ip) and Unemployment Rate Potentials (Up)) are calculated from the observed values of the variables. This simply means that a transformation is made to each variable in order to insert the spatial relationships among elements of each of the variable. And so the regression model changed to a form that the explanatory variables are measured on the average with spatial interactions. The transformation might be said to be linear

[1]In his paper, a detailed derivation of relative income potential was presented, which is not repeated in this text.

to each element among variables. This produces a high correlation among variables. In Table 3.1, a correlation matrix among variables is provided for the comparison of the correlation changes before and after the transformation.

The potentials' regression is hypothesized to give a better result because the spatial interaction effects of in-migration have been accounted for by using potential regressors. If the potentials really can explain more about the variation of per capita in-migration, i.e., in-migration divided by the total population (M/P), then a further hypothesis can be made to the predicted in-migration converted from the predicted M/P by using potential regressors would be more efficient than that obtained from observed variables regression. These hypotheses will be tested in the following chapters.

3.3 The Potential

The potentials calculated through out this study use the following formula:

$$ {}_iV = \sum_{j=1}^{n} \frac{V_j}{D_{ij}} \tag{3.2} $$

where, D_{ij} = highway mileage between region i and j;

${}_iV$ = total base potential of region i;

V_j = mass in region j; and

$j = 1, 2, ..., n$.

Since we needed to provide the cohesive force of the region from the population within the region, we could not neither simply ignore the own distance nor just set it to zero (If it is set to zero, the potential will become infinity.). So, many forms appeared to determine the self-potential of regions. Carrothers used the deviation of potential created by region j at region i as an approximation of i's self-potential. Another common approximation of self-potential was to pick the own distance measure as the distance from the region's center to its periphery (see Stewart [18]). The above examples are only the ones of many self-potential definitions used in different studies. It is important to notice that there is no such an unique method to approximate the self-potential since it is really hard to be measured. How about using the radius from the center to the region boundary as own distance measure? How about using diameter? How about using areas ratios? How about using masses ratios? Different researchers might take different approaches and come up with different self-potentials. In this study, we use half of the longest diagonal of the region for the own distance measures; see Appendix E.

Another kind of potential measure is called base potential, which is introduced by Stewart [20]. He used the base potential to produce population potential in his population distribution model. The base potential excludes the self-potential of a region. In this study, we cannot ignore self-potential. Although in-migration within the own region is not defined as in-migration, but there exists a self attractional force preventing the individual from leaving or moving out of the region. Or, may be we can loosely call this the opportunity cost of moving. That is why we use the common potential definition.

Table 3.1: Correlation Matrix of Variables and Potentials

	LNMP87	LNMP76	P70	P60	U70	U60	I70	I60	Pp70	Pp60	Up70	Up60	Ip70	Ip60
LNMP87	1.00	.895	-.420	-.449	.168	-.040	.115	.005	-.539	-.551	-.504	-.555	-.521	-.531
LNMP76		1.00	-.326	-.371	.145	-.094	.319	.187	-.365	-.380	-.333	-.398	-.348	-.355
P70			1.00	.994	-.052	-.043	.183	.192	.267	.257	.115	.122	.140	.143
P60				1.00	-.049	-.029	.124	.144	.269	.261	.125	.135	.150	.153
U70					1.00	.714	-.090	-.007	.376	-.382	-.274	-.364	-.429	-.407
U60						1.00	-.179	-.108	-.217	-.216	-.205	-.174	-.298	-.284
I70							1.00	.897	.250	.232	.229	.169	.237	.245
I60								1.00	.265	.250	.251	.183	.242	.268
Pp70									1.00	.999	.848	.892	.905	.907
Pp60										1.00	.849	.897	.909	.911
Up70											1.00	.969	.967	.968
Up60												1.00	.972	.968
Ip70													1.00	.998
Ip60														1.00

4. THE DATA AND THE POTENTIALS

4.1 Data Definition

In order to study in-migration flows for different time spans, we need to freeze the spatial dimension. Data used in this study are based on definitions of the Control Regions (CRs). The definition of the CRs are taken according to the Metropolitan Statistical Areas (MSAs) boundaries published in the County and City Data Book (C.C.D.B.) [21] which defined MSAs on April, 1984. Having sub-divided the Consolidated MSAs into some smaller Primary MSAs, a total of 313 MSAs are defined as Metropolitan Control Regions(MCRs). The rural area of each state which does not classify as MSA is classified as Non-metropolitan Control Region (NCRs). Therefore, there are 49 NCRs in the U.S. (New Jersey state and Washington D.C. have no NCR), and altogether 362 CRs. A list of the CRs is printed in Appendix A. Because of the re-definition of regions, the data are obtained by regrouping and recalculating from county's data. Therefore, Census Reports and the C.C.D.B. are the only sources for data collection.

4.1.1 The Dependent Variable

4.1.1.1 Per Capita In-migration[1] ($M(t)/P(t')$): Per capita in-migration is obtained by dividing the in-migration of period t by the population of period t' (= t - 10). That is, if t denotes the period of years 1965 - 1970, we call t = 1970, and then, t' = 1960. There are two reasons to make this expression. First, to consider the causality effect, it is more meaningful to study the in-migration to past population, P(t'), than that to the present population, P(t). It is because the former can tell how future in-migration is related to the present population. Secondly, to consider the prediction of in-migration, P(t) can then be used to predict the in-migration at time t + 10. That is, a more certain information can be used in the prediction procedure. In-migration is defined as the internal population shifts across the CRs within the U.S. So, the data under the title "Change of Residence" in the Census Reports [22] are collected. Under this title, there are two items provided the data to the in-migration variable. One item is the number of people who moved to county i from different counties in the same state within 5 years before the census is made. Another item is the number of people who moved to county i from different counties from different states within 5 years before the census is made. Therefore, when in-migration of a multi-county CR is collected, problems arise if the totals of in-migrants from the same state or different states of every county within the CR are summed up, because the sum is now involved not just with the inter-CR migrants but also the intra-CR-inter-county migrants. Some in-migrants who are not defined as the

[1] Per Capita In-migration in the regressions studied and in this text is referred to the In-migration at time t, M(t), over the lagged Population, P(t - 10), e.g., M(80)/P(70) means that in-migration rate of the CR during the period 1975 to 1980 in terms of the population in 1970, we call it MP87.

in-migrants to a CR but are counted as if they were. A weighting method applied to net out these intra-CR-inter-county in-migrants will be discussed in the next section.

4.1.2 The Independent Variables

4.1.2.1 Population: Population of each CR is the total population of the county or the sum of the total population of the counties in the CR. See the column "Population" by county in the C.C.D.B.

4.1.2.2 Per Capita Real Income: There are at least three types of income used in the past studies. Carrothers used "Personal Income", while Schwartz [17] used "Median Income". One income definition used in this study is the Per Capita Income because it measures the average income for each interactor in each CR. It also measures the attraction of the probability of expected income of moving. For the need of comparing the standard of living of the interactors over time and across CRs, the deflated real income is applied. The Per Capita Real Income is obtained from deflating the current per capita income by the Consumer Price Index for all Urban consumers (CPI(U)) and also the Regional Price Index. The annual average indices of 4 regions across 4 city sizes are used to distinguish the different real income levels of the CRs. An assumption of the fixed price level pattern over time is assumed. A table of the CPI(U) is printed in Appendix B.

To generate the Regional Price Index, we assumed the consumption pattern is fixed for each individual, so the prices of different commodities across regions are proportional to the expenditures of on those commodities. We use the expenditures on the 7 commodities[2] of the rural region as the base value to approximate the relative regional prices. The log of the relative prices are then regressed on the regional dummy variables, so that the exponential of the coefficients can be used as the regional price of each commodity. Then, we can multiply the approximated prices with the relative importance to form the regional price index; see Appendix B.

4.1.2.3 Unemployment Rate: Unemployment Rate is the rate of the number of unemployed civilian labor force divided by the total number of civilian labor force in the CR.

4.1.2.4 Distance: Distance measure had been discussed in many literatures. A common use of distance measure is the highway distance measure, for example, Anderson [1]. He suggested that,

> " ... highway mileage was the best available measure of "shortest transportation distance". ... "

In this study, 301 MCRs and 49 NCRs centers can be found directly on the Standard Highway Mileage Guide (S.H.M.G.). But 12 CRs[3], say CR_m, which do not exist on the S.H.M.G. have to measure on the Road Altas the shortest highway mileage to the nearest city, the 'Link City'. The link cities are the cities which can be found on the S.H.M.G. Then, the highway mileage of CR_m to other CRs are calculated by adjusting the distance to each of the other CRs to the distance

[2] They are food, alcoholic beverage, housing, apparel and services, transportation, health and entertainment. The relative importance of components of CPI in 1986 is used in the calculation of the Regional Price Index.

[3] 11 MCRs and 1 NCR, including Hawaii.

between CR_m and its link city; see Appendix C. The distance conditions among CRs are assumed to be fixed over time. The intra-CR-inter-county distances printed in Appendix D are the direct distances measured by hand with a ruler on the state maps in the C.C.D.B. The measures are double checked to minimize the incorrect measurement.

4.1.3 In-migration Estimate

In-migration of the CR_k can be estimated from its counties' totals of in-migration from other counties in the same state, M_k^s and that from the other counties in different states, M_k^d. The in-migration total of each of the county can be calculated by adding up the in-migration to the county from different counties in the same state and the in-migration from different counties in different states. Since the total of counties' in-migration incurred intra-CR-inter-county (ICRIC) in-migration, a weighting procedure is applied to net out these ICRIC counts. The weight is based on the inverse relationship of in-migration and distance proposed by the gravity theory in multi-regional framework. There are three kinds of structures a MCR can exist. First, a Mono-county-CR consists of one and only one county. It is the most simple form because the in-migration can be calculated by adding up the migrant total from different counties of same state and from different states, see Figure 4.1. Secondly, a Multi-county-CR consists of more than one county and is located within a particular state boundary. This kind of structure causes trouble on the M_k^s, see Figure 4.2. That is, when the counties' in-migration is summed up, the in-migration from b and c to a, from a and c to b and from a and b to c are also counted. However, these in-migration numbers will be the ones needed to be excluded.

The third kind of structure can be subdivided into three kinds of structures as shown in Figures 4.3, 4.4 and 4.5, which are called the Multi-county-multi-state-CRs (MCMS-CRs). They consist of at least two counties and locate across the state boundaries. The examples shown in the following figures are only the most simple 2, 3 and 4-county-multi-state-CRs. The CR structure in Figure 4.3 causes trouble on both M_k^s and M_k^d. For the problem on the same state total, county c' to county a' in-migration is counted; for the problem on different state, counties b' and d' to a' in-migrations are counted. Similar errors occurred to counties b', c' and d'. So, both the M_k^d and M_k^s need to be adjusted.

Another CR structure under the same title consists of two counties and each of the two locates in a separate state, as shown in Figure 4.4. This CR causes trouble on M_k^d but not on M_k^s.

The third Multi-county-multi-state-CR structure is shown in Figure 4.5, which is a combination of the other two structures under the same title. It causes problem on different state total at county a'' because counties b'' and c'' in-migrations to a'' are also counted, it has no trouble on the same state at a''. But it has trouble on both different state total and the same state total at b'' and c''.

Because of the differentiations among structures, different weighting strategies are applied. According to Foot and Milne [6] a weighting method can be in the form,

$$W_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j}^{n} \frac{1}{d_{ij}}} \tag{4.1}$$

where, W_{ij} = weight for interactions for j to i;

i = origin city;

j = interaction city; and $i \neq j$.

Modified forms of the above weight are made to adjust for different CR structures. It is easily to observe that the Structure 1 of the MCMS-CR_k can be a control case of all other structures, i.e., all other structures are only its special cases. To solve the problem, we first construct the following formula to weight out the double counted in-migrant, for MCMS-CR_k (Structure 1):

$$M_k = \sum_{i=1}^{I_k} \left[M_i^s \Big(1 - \frac{\sum_{j=1}^{J} \frac{1}{d_{ij}}}{\sum_{j=1}^{J} \frac{1}{d_{ij}} + \sum_{k'=1}^{K'} \frac{1}{d_{kk'}}} \Big) + M_i^d \Big(1 - \frac{\sum_{p=1}^{P} \frac{1}{d_{ip}}}{\sum_{p=1}^{P} \frac{1}{d_{ip}} + \sum_{k''=1}^{K''} \frac{1}{d_{kk''}}} \Big) \right] \tag{4.2}$$

where, k = particular MCR in study, k=1, 2, ..., 313;

M_k = total internal in-migration to MCR_k;

i = county within MCR_k;

I_k = number of county(ies) MCR_k contains;

M_i^s = total internal in-migration to the county i of MCR_k from the same state as county i located;

M_i^d = total internal in-migration to the county i of MCR_k from states other than county i located;

d_{ab} = distance between the regions a and b, (a, b) can be counties or CRs depends on the subject to be measured;

j = other county(ies) within the same MCR in the same state;

k' = other CRs (which have at least one county) located in the same state;

p = other county(ies) within the same MCR in the different state;

k'' = other CRs (which have at least one county) located in the different state;

J = Maximum number of j;

P = Maximum number of p;

K' = Maximum number of k'; and

K'' = Maximum number of k''.

For the Mono-county-CR:

$$\sum_{j=1}^{J} \frac{1}{d_{ij}} = 0$$

and;

$$\sum_{p=1}^{P} \frac{1}{d_{ip}} = 0$$

then,

$$M_k = M_i = M_i^s + M_i^d.$$

For the Multi-county-CR:

$$\sum_{p=1}^{P} \frac{1}{d_{ip}} = 0$$

then,

$$M_k = \sum_{i=1}^{I_k} \left[M_i^s \Bigg(1 - \frac{\sum_{j=1}^{J} \frac{1}{d_{ij}}}{\sum_{j=1}^{J} \frac{1}{d_{ij}} + \sum_{k'=1}^{K'} \frac{1}{d_{kk'}}} \Bigg) + M_i^d \right].$$

For MCMS-CR (Structure 2):

$$\sum_{j=1}^{J} \frac{1}{d_{ij}} = 0$$

then,

$$M_k = \sum_{i=1}^{I_k} \left[M_i^s + M_i^d \Big(1 - \frac{\sum_{p=1}^{P} \frac{1}{d_{ip}}}{\sum_{p=1}^{P} \frac{1}{d_{ip}} + \sum_{k''=1}^{K''} \frac{1}{d_{kk''}}} \Big) \right].$$

For MCMS-CR (Structure 3):

$$\sum_{j=1}^{J} \frac{1}{d_{i'j}} = 0$$

then,

$$\begin{aligned} M_k \;=\; & \sum_{i'=1}^{I'_k} M_i'^s + \sum_{i''=1}^{I''_k} \left[M_i''^s \Big(1 - \frac{\sum_{j=1}^{J} \frac{1}{d_{i''j}}}{\sum_{j=1}^{J} \frac{1}{d_{i''j}} + \sum_{k'=1}^{K'} \frac{1}{d_{kk'}}} \Big) \right. \\ & \left. + M_i^d \Big(1 - \frac{\sum_{p=1}^{P} \frac{1}{d_{ip}}}{\sum_{p=1}^{P} \frac{1}{d_{ip}} + \sum_{k''=1}^{K''} \frac{1}{d_{kk''}}} \Big) \right] \end{aligned} \tag{4.3}$$

where, I'_k = Maximum number of i' in MCR_k;

I''_k = Maximum number of i'' in MCR_k;

$I_k = I'_k + I''_k$;

i' = county of Multi-county-MCR_k stands alone in a state;

i'' = county of Multi-county-MCR_k which has at least one other county from the same state.

According to the five strategies listed above, the estimated in-migrations of the MCRs are obtained. To compare the estimated in-migration with the observed values, 160 out of 313 MCRs in 1980 are Mono-county-CRs, so the estimated in-migration and the observed in-migration are exactly the same; 72 Multi-county or Multi-county-multi-state-CRs are the same defined as the census regions and therefore the estimated values, $\widehat{M80}$, and the observed values, M80, can be compared; see Table 4.1.

A graph showing the relationship between the two values is plotted in Figure 4.6. The observed values are regressed on the estimated values by using SAS PROC REG procedure with no intercept option. Therefore, the regression line starts from the origin. If the estimated values and the corresponding observed values are all the identical, then the coefficient, the slope of the regression, is expected to be 1.00. Since we picked the observed values as the standard values in this comparison, the coefficient obtained is 0.95 that means the estimated values are on the average 5 % underestimated the observed values from the census report in 1980.

Note that the above estimations are made for only the MCRs, how about the NCRs? The estimation for the NCR_l is obtained by using the state total in-migration from same state (M_l^s) and from different states (M_l^d). As shown in Figure 4.7, the in-migrant of NCR_l is supplied by in-migrants from different state and in-migrants from the MCRs within the same state l.

The in-migrant from different state ($\tilde{M}_l^d$) can be obtained by subtracting the different states' total in-migration of all MCR counties from the M_l^d. That is,

$$\tilde{M}_l^d = M_l^d - \sum_{k=1}^{K_l} M_k^d \tag{4.4}$$

where, K_l = number of MCR counties in state l.

For the same state in-migration total, a simpler weighting method is chosen, otherwise, there will be too many inter-county distance measures for the NCRs. For instance, there are 20,910 distance measures for NCR of TEXAS, 4,753 for that of MISSOURI and etc. So, number of county ratios will be used as the weights,

$$\tilde{M}_l^s = M_l^s \cdot [\frac{TTLMCR}{(TTLSTATE)-1}]$$

where, TTLMCR = Total number of counties of MCRs in state l; and

TTLSTATE = Total number of counties in state l.

Since M_l^s is the county's in-migration total of the state l and the in-migration to NCR_l from same state must come from the MCRs counties in the same state (minus one because in-migrations count from other counties). And finally, the NCR_l in-migration can be obtained from:

$$M_{(NCR_l)} = \tilde{M}_l^d + \tilde{M}_l^s.$$

4.1.4 Potential Formation

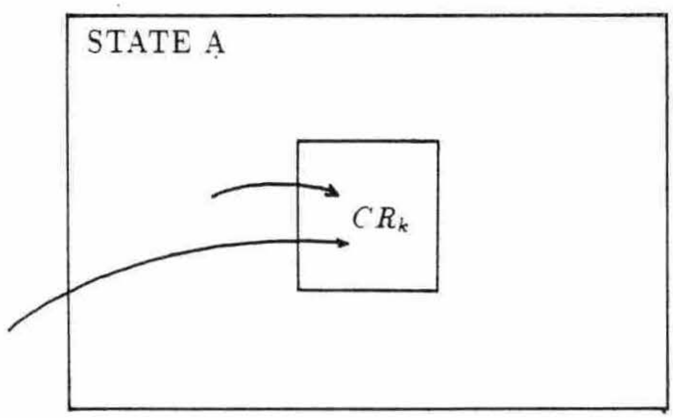
To calculate the potentials from the observed values of the variables, the following formula is used.

$$_iX_p = \sum_{j=1} \frac{X_j}{d_{ij}} \qquad (4.5)$$

where, $_iX_p$ = total potential of X at CR_i; and

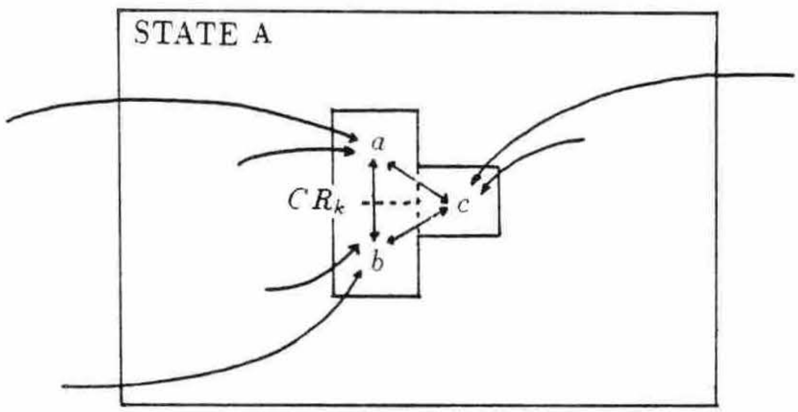
X_j = observed value of X at CR_j; j= 1,2,...,J(=362).

By substituting X with the observed values P, I and U, the Population Potential (Pp), Income Potential (Ip) and Unemployment rate Potential (Up) are obtained. Through this transformation from observed values to potentials, the spatial interaction of each of the variables is inserted. But one thing that should be noticed is that the potentials are very sensitive to distance among the CRs. That is, they are very dependent on how the CRs are defined. In other words, a CR with high or low observed values may or may not have a high or low potential, but the insertion of spatial relationship to variables creates an on-the-average relationships for each potential measures.



$\longrightarrow$ = In-migration from other CRs.

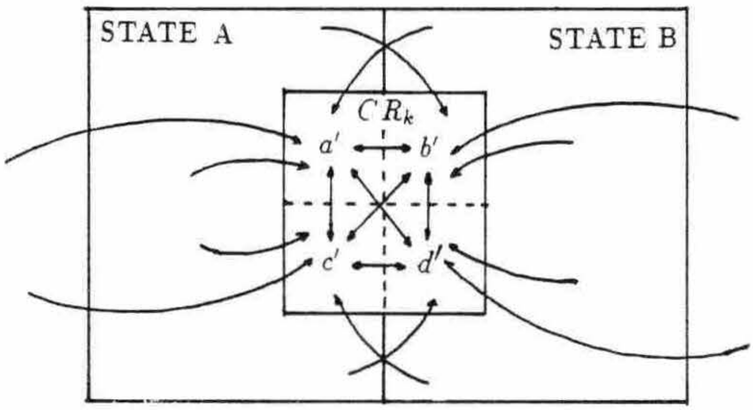
Figure 4.1: A Mono-county-CR



$\longrightarrow$ = In-migration from other CRs.

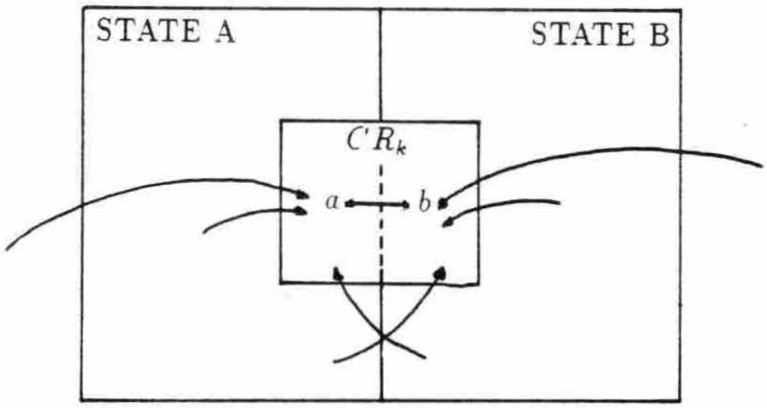
$\longleftrightarrow$ = Intra-CR-inter-county In-migration.

Figure 4.2: A Multi-county-CR



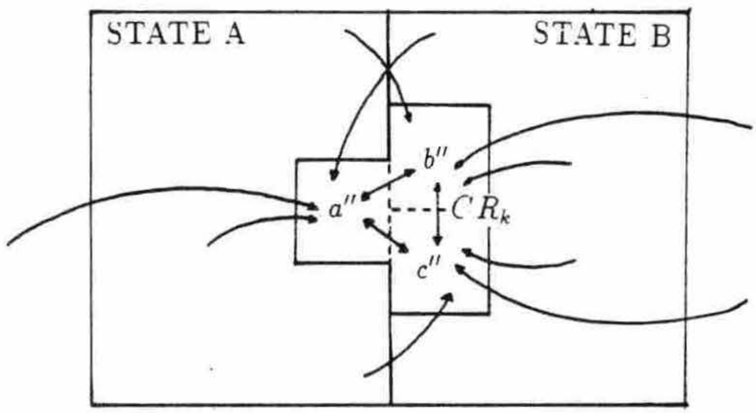
$\longrightarrow$ = In-migration from other CRs.
$\longleftrightarrow$ = Intra-CR-inter-county In-migration.

Figure 4.3: MCMS-CR (Structure 1)



$\longrightarrow$ = In-migration from other CRs.
$\longleftrightarrow$ = Intra-CR-inter-county In-migration.

Figure 4.4: MCMS-CR (Structure 2)



$\longrightarrow$ = In-migration from other CRs.
$\longleftrightarrow$ = Intra-CR-inter-county In-migration.

Figure 4.5: MCMS-CR (Structure 3)

Table 4.1: Estimated Migration for 72 CRs in 1980

CR	M80	$\widehat{M80}$	CR	M80	$\widehat{M80}$
2	78120	79477	147	77844	77070
3	22284	19216	149	10372	9577
7	66771	73147	164	63114	43726
9	40190	42454	168	34064	33455
16	37759	36181	181	115986	108777
18	35870	20780	185	120271	144723
23	152853	139514	187	58359	55692
27	75511	59684	191	46270	42988
29	46178	44768	195	120366	107664
38	17171	12693	196	260580	268180
53	40181	38788	212	191063	194534
58	29547	29568	218	63761	63803
60	32350	24460	219	41853	44484
61	52479	47437	220	324660	402695
64	129688	130403	233	102294	102394
65	40647	40094	237	47893	45103
66	135227	155569	239	437942	443942
67	105093	102800	242	94410	77645
69	86846	91168	243	31173	30488
70	52192	48659	246	31587	21149
72	59233	52860	249	70468	63621
73	10145	11457	253	171507	153594
75	9893	12384	264	282207	319505
76	47500	49432	269	15593	16828
85	34054	35224	273	27836	24156
86	25052	22492	274	45948	37895
94	33404	37317	277	12551	12984
98	15451	14149	279	76537	70905
112	54018	62735	292	30264	28952
113	15348	15413	293	95411	92086
115	72541	77016	303	14428	16566
120	79278	74962	304	70258	59620
131	130344	130163	307	64276	62144
139	48031	37759	311	44969	46273
140	16498	18266	312	35307	39458
142	22776	23403	313	25899	25876

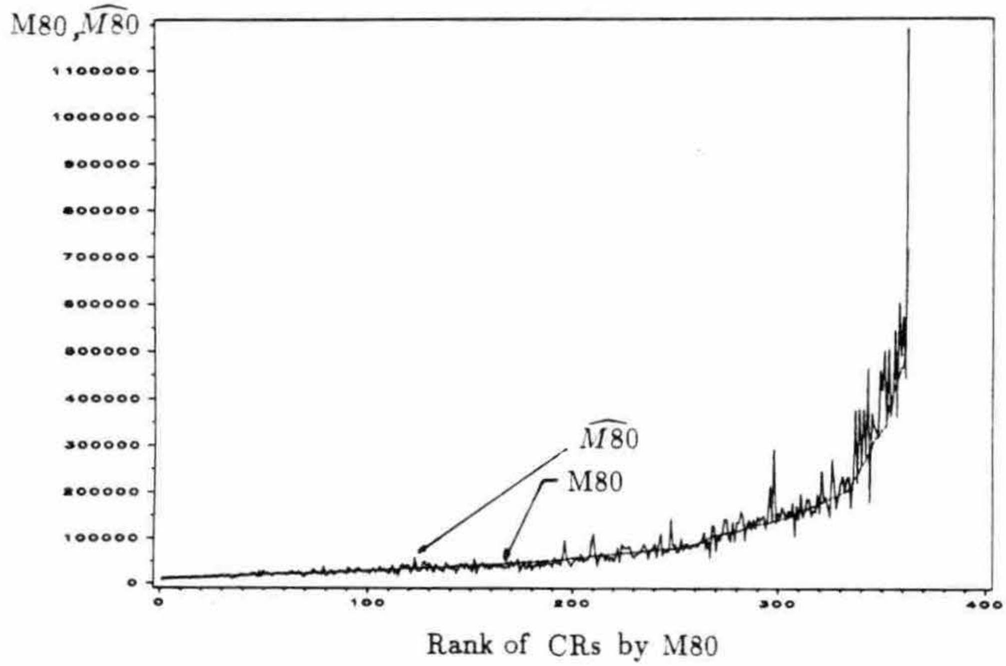


Figure 4.6: Est. Vs. Observed Migration in 1980

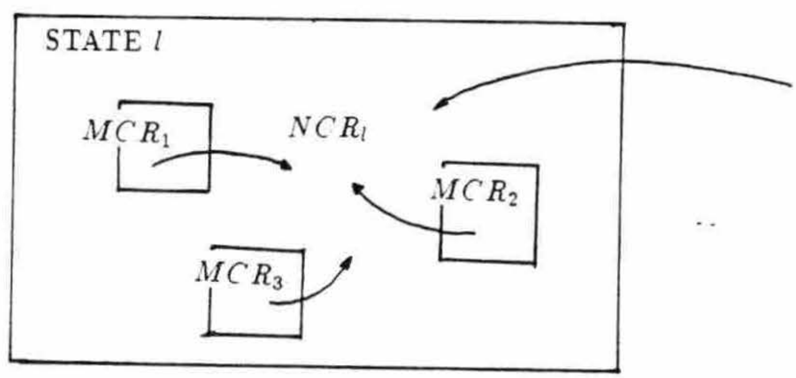


Figure 4.7: Migration of NCR_l

4.2 An Illustration of Distance Effects

The following is a simple illustration for the four regions example to show how the transformation affects the regional relationship. In Figure 4.8, 4 CRs are analyzed and their inter-regional distances are linked among the centers.

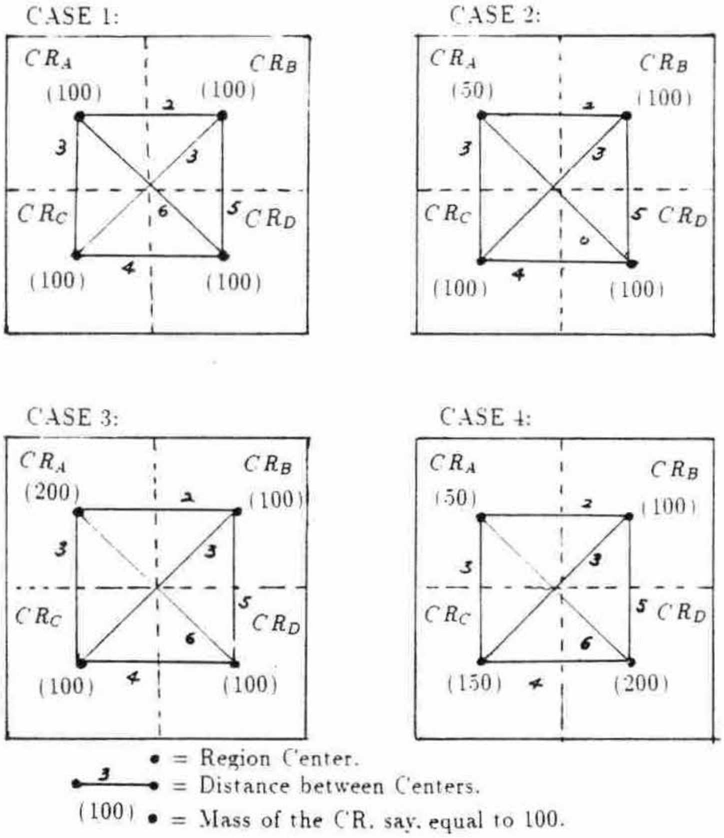


Figure 4.8: Illustration: Total Potential of 4 CRs Cases

Then, from the Figure 4.8, the following results can be obtained (Note: Assume the own distance equal to 4 for all regions, and the mass of each region is marked on the figure.).

Case 1: $A = B = C = D$

Total Potential of A (Tpot A) = 100/2 + 100/3 + 100/6 + 100/4= 125

Total Potential of B (Tpot B) = 100/2 + 100/3 + 100/5 + 100/4 = 128.3

Similarly, the Tpot C and Tpot D are,

Tpot C = 116.7
Tpot D = 86.7
Result 1: $TpotB > TpotA > TpotC > TpotD$.

Case 2: $B = C = D > A$

Tpot A = 112.5
Tpot B = 103.3
Tpot C = 100
Tpot D = 78.3
Result 2: $TpotA > TpotB > TpotC > TpotD$.

Case 3: $A > B = C = D$

Tpot A = 150
Tpot B = 178.3
Tpot C = 150
Tpot D = 103.3
Result 3: $TpotB > TpotC = TpotA > TpotD$.

Case 4: $D > C > B > A$

Tpot A = 112.5
Tpot B = 140
Tpot C = 137.5
Tpot D = 115.8
Result 4: $TpotB > TpotC > TpotD > TpotA$.

From the results above, although the total potential of D is the smallest in all four cases, but the rank of the total potential of A, B and C cannot be determined or related to their observed values. This is because the potentials are very sensitive to the distance structure, or simply say, their locations. Therefore, the relationship between the observed values and their potentials are highly dependent on the distribution of the observations and the overall distance structure.

5. FUNCTIONAL FORMS AND MODELS

5.1 Functional Forms

Considering the final goal of this study is to predict the future in-migration, the natural log of the per capita in-migration is used as the dependent variable in order to guarantee obtaining positive predicted values. On the other hand, two common functional forms are chosen to explain the per capita in-migration by potentials. Those forms are,

(1) Exponential Equational Form:

Basic structure:

$$Y = a_o' \cdot \exp(a_1 \cdot X_1 + a_2 \cdot X_2 + \cdots + a_n \cdot X_n) \cdot V'.$$

Transformed structure:

$$\ln Y = a_o + a_1 \cdot X_1 + a_2 \cdot X_2 + \cdots + a_n \cdot X_n + V.$$

(2) The Cobb-Douglas Equational Form:

Basic structure:

$$Y = a_o' \cdot X_1^{a1} \cdot X_2^{a2} \cdots X_n^{an} \cdot V'.$$

Transformed structure:

$$\ln Y = a_o + a_1 \cdot \ln X_1 + a_2 \cdot \ln X_2 + \cdots + a_n \cdot \ln X_n + V.$$

These forms are chosen in this study because they are simple to explain the relationship of variables in both their basic and transformed structures. The dependent variable is a linear combination of the independent variables in the transformed form. This makes it easier to be applied linear regression analysis and to interpret the results.

Two regressions of 724 (=2x362) observations each are run to help estimating a functional form that will be used in the model building procedure. The first regression (Reg 1) regressed a pooled M(80)/P(70) and M(70)/P(60) on the pooled independent variables of 1980's and 1970's. If, in the general pattern, an assumption of the responsive lag is made, then another regression (Reg 2) is run. It has the same structure in the dependent variable, but is regressed on the pooled lagged independent variables of the 70's and 60's. It is used to check whether the regression on the lagged independent variables is significantly different, or better fitted, from using the unlagged independent variables, and to determine whether the responsive lag is an appropriate assumption.

The results of regressing the selected functional forms are shown in Table 5.1:

Table 5.1: Functional Forms Regression Results

Functional Form	Reg	R^2
Exponential	1	0.8450
	2	0.7605
Cobb-Douglas	1	0.9201
	2	0.9351

According to the pooled data regression, Cobb-Douglas Equational Form, has $R^2 > 0.9$ in both cases, suggested to be the better form whether or not the responsive lag assumption is presented.

On the other hand, the Box-Cox Regression is run to help identifying the functional form to be used. The Box-Cox Transformation equation has the form,

$$\frac{Y^{\lambda_o}-1}{\lambda_o} = a_o + a_1\frac{X_1^{\lambda_1}-1}{\lambda_1} + a_2\frac{X_2^{\lambda_2}-1}{\lambda_2} + \cdots + a_n\frac{X_n^{\lambda_n}-1}{\lambda_n} + U.$$

Box-Cox Regression suggests that if $\lambda_i = 0$; $i = 0,1,2,\cdots,n$; then $\ln Y$ and $\ln X_j$, j=1,2,$\cdots$,n, will be used in the model. If $\lambda_i \neq 0$; i=0,1,2, $\cdots$,n; then $\frac{Y^{\lambda_o-1}}{\lambda_o}$ and $\frac{X_j^{\lambda_j}}{\lambda_j}$, j=1,2,$\cdots$,n, will be used in the model.

The Box-Cox Regressions are run by regressing M(80)/P(70) on independent variables of the 70's (model 1) and the M(70)/P(60) on the 60's (model 2). The regression can double check whether the general functional form of Cobb-Douglas equation is the correct form to be specified. By restricting all λ_j's are equal, i.e., $\lambda_o = \lambda_i = \lambda$, the same transformation will be done to all variables in the model if needed. Then, a comparison of Cobb-Douglas equational form and linear form are made by restricting the λ_i's to 0 and 1, respectively. The results are shown in Table 5.2.

The Box-Cox Regression suggests that a log transformation to the variables is appropriate since, in the first restriction on the Table 5.2, the λ obtained in both models are -0.26 and -0.11 (both close to zero). On the other hand, the log transformation ($\lambda_o = \lambda_1 = \cdots = \lambda_n = 0$) models are seemed to be superior to the linear ($\lambda_0 = \lambda_1 = \cdots = \lambda_n = 1$) models. The R^2 of the log models are 0.2986 and 0.1849 which are higher than that of without transformation restriction.

The log transformations of variables can not only be a better model of the relationship between the per capita in-migration and the explanatory variables, but also, in most cases, stabilize the variances of the variables.

Table 5.2: Box-Cox Regression Results

Restrictions	Model	λ	R^2	LLF
$\lambda_0 = \lambda_1 = \cdots = \lambda_n = \lambda$	1	-0.11	0.3092	378.9
	2	-0.26	0.1822	391.5
$\lambda_0 = \lambda_1 = \cdots = \lambda_n = 0$	1	0.00	0.2986	374.9
	2	0.00	0.1849	386.2
$\lambda_0 = \lambda_1 = \cdots = \lambda_n = 1$	1	1.00	0.2524	286.2
	2	1.00	0.1525	268.6

LLF=Log-Likelihood Function.

5.2 Models and Ideas of Model Shifting

From the above section, the chosen model has the BASIC form:

$$LNMP(t) = a_o + a_1 \cdot LNPp(t) + a_2 \cdot LNUp(t) + a_3 \cdot LNIp(t) + V_t \qquad (5.1)$$

where, LNMP(t) = natural log of $M(t)/P(t')$ ratio at year t;

LNPp(t) = natural log of Population Potential in the year t;

LNUp(t) = natural log of Unemployment rate Potential in the year t;

LNIp(t) = natural log of per capita real Income Potential in the year t;

V_t = disturbance term;

t = 70, 80; and t'= t - 10.

The next question is how to modify from this BASIC model so that a better regression model can be obtained? Six modified models are discussed below.

5.2.1 The Responsive Lag (RL) model

The responsive lag assumes that in-migration behaviors of the present time depended on the observed information in the past. This is a reasonable assumption to the general public. Information come from relatives, friends and published statistics are considered before a person makes decision to migrate. In this study, because of restrictions on data sources, the shortest length of time allowed for the migration response is 10 years, represented by "t'". Then, the RL model has the form:

$$LNMP(t) = a_o + a_1 \cdot LNPp(t') + a_2 \cdot LNUp(t') + a_3 \cdot LNIp(t') + V_t. \quad (5.2)$$

where t = 70 and 80 for LNMP(t); also see labels of equation (5.1).

5.2.2 The Responsive Lag-Lag (RLL) model

The RLL model follows the RL model and further assumes the in-migration population ratio or the per capita in-migration dependent on its past performance. This assumption is reasonable since the in-migrants moved to a place affect the future in-migration to that place. The activities made by the earlier in-migrants can alter the attractional force of that place, then they influence the in-migration in the following years. So, an additional variable, the lagged per capita in-migration ($LNMP(t')$), is introduced to the RL model. The structural form of the RLL model is:

$$\begin{aligned} LNMP(t) &= a_O + a_1 \cdot LNPp(t') + a_2 \cdot LNUp(t') + a_3 \cdot LNIp(t') \\ &\quad + a_4 \cdot LNMP(t') + V_t. \end{aligned} \tag{5.3}$$

Problems raised for the RLL model because the model looks like autoregressive to the dependent variable. On the one hand, the lagged per capita in-migration is expected to capture most of the information and dominate the importance in the analysis. The advantage is to obtain high explanatory ability. The disadvantage is the ignorance of some other significant explanatory variables by employing the lagged dependent variable. On the other hand, this model may not obtain minimized variance so that the estimators are not BLUE (Best Linear Unbiased Estimator) within the linear estimator family. Therefore, two methods, by differencing and by substituting the dependent variable, are used to try to eliminate or reduce the autoregressiveness in the model.

5.2.3 The Difference (DIFF) model

One method to deal with the autoregression problem is to use the differential term of the per capita in-migration, $LNMP(t) - LNMP(t')$, substitute for the lagged per capita in-migration as the explanatory variable. According to Kau and Sirmans [12], this method can reduce the dominant of the lagged term and keep the significance of other variables. The structural form of the DIFF model is

$$\begin{aligned} LNMP(t) &= a_o + a_1 \cdot LNPp(t') + a_2 \cdot LNUp(t') + a_3 \cdot LNIp(t') \\ &\quad + a_4 \cdot LNMPD(t) + V(t) \end{aligned} \tag{5.4}$$

where, $LNMPD(t) = LNMP(t) - LNMP(t')$.

5.2.4 The Substitution (SUBT) model

Since the nature of the dataset of 362 observations of each of the time span and the restriction for using only Ip, Up and Pp as regressors, we have no intention to put in some other variables as instruments. So, the Instrumental Variable method suggested in Judge, Griffiths, Lütkepohl, and Lee [10] to solve autoregression problem will not be applied to the RLL model. A substitution method is used. This method depends highly on the consistent relationship between the dependent and independent variables. In the RL model,

$$LNMP87 = f_1(LNPp70, LNUp70, LNIp70) \tag{5.5}$$

and,

$$LNMP76 = f_2(LNPp60, LNUp60, LNIp60) \tag{5.6}$$

so if,

$$LNMP87 = g(LNMP76, LNPp70, LNUp70, LNIp70) \tag{5.7}$$

then,

$$\begin{aligned} LNMP87 &= h(LNPp60, LNUp60, LNIp60, LNPp70, \\ &\quad LNUp70, LNIp70). \end{aligned} \tag{5.8}$$

This substitution is good to keep the generality of the variable, and the model is in the form,

$$\begin{aligned} LNMP87 &= a_o + a_1 \cdot LNPp60 + a_2 \cdot LNUp60 + a_3 \cdot LNIp60 \\ &\quad + a_4 \cdot LNPp70 + a_5 \cdot LNUp70 \\ &\quad + a_6 \cdot LNIp70 + V. \end{aligned} \tag{5.9}$$

The disadvantage of this model is that it generates of high multi-collinearity problem among the regressors.

5.2.5 The Residual (RES) model

One thought is to study the residuals produced in the past regression in the current regression model. Since from RLL model, the $V_{t'}$ is correlated with $LNMP(t)$ and so V_t. A model including $V_{t'}$ as regressor may be an alternative to account for the rest of the effects other than the random effect. To consider whether to put $V_{t'}$ as one of the regressors, if the RL model is not a perfect model of the realization, i.e., some significant variables may have been excluded from the model, those excluded but valuable sources of information are mixed with the true random errors. This can be seen by a low R^2 of the RL model. So, if the residual of 70's RL model is included into the regression RL model of 80's, it becomes the RES model, the R^2 as well as explanatory ability are expected to increase. The structural form of model will be:

$$\begin{aligned} LNMP87 &= a_o + a_1 \cdot LNPp70 + a_2 \cdot LNUp70 + a_3 \cdot LNIp70 \\ &\quad + a_4 \cdot RES76 + V. \end{aligned} \tag{5.10}$$

where, RES76 = LNMP76 - $LN\widehat{MP}76$;

and,

$$LNMP76 = b_o + b_1 \cdot LNPp60 + b_2 \cdot LNUp60 + b_3 \cdot LNIp60 + V'. \tag{5.11}$$

Although RES76 is a function for LNMP76 and some explanatory variables of the 60's, RES76 can be seen as an independent variable because it is obtained only after the 70's RL regression is done and is given unchanged in the 80's regression model. One disadvantage of this model is that the RES76 again provided high multicollinearity with other regressors.

5.2.6 The AutoRegressive (AR) model

Having changed a little bit from the RES model, an assumption is made by assuming the residuals are in first order autocorrelation form, i.e., $V_t = \rho \cdot V_{t-1} + e_t$, $e_t \sim i.i.d.(0, \sigma_e^2)$ in the RL model, AutoRegressive model of first order, or AR(1) model, can be applied. The model assumes,

$$Y_t = a_o + a_1 \cdot X_{1,t-1} + a_2 \cdot X_{2,t-1} + a_3 \cdot X_{3,t-1} + V_t \tag{5.12}$$

where, $V_t = \rho \cdot V_{t-1} + e_t$; and $e_t \sim i.i.d.(0, \sigma_e^2)$

then,

$$\rho \cdot Y_{t-1} = \rho \cdot a_o + \rho \cdot a_1 \cdot X_{1,t-2} + \rho \cdot a_2 \cdot X_{2,t-2} + \rho \cdot a_3 \cdot X_{3,t-2} + \rho \cdot V_{t-1}$$

$$\begin{aligned} Y_t - \rho \cdot Y_{t-1} &= a_o \cdot (1-\rho) + a_1 \cdot (X_{1,t-1} - \rho \cdot X_{1,t-2}) \\ &\quad + a_2 \cdot (X_{2,t-1} - \rho \cdot X_{2,t-2}) + a_3 \cdot (X_{1,t-1} \\ &\quad - \rho \cdot X_{3,t-2}) + e_t. \end{aligned} \tag{5.13}$$

Therefore,

$$\tilde{Y}_t = \tilde{a}_o + a_1 \cdot \tilde{X}_{1,t-1} + a_2 \cdot \tilde{X}_{2,t-1} + a_3 \cdot \tilde{X}_{3,t-1} + e_t$$

where, $\tilde{Y}_t = Y_t - \rho \cdot Y_{t-1}$;

$\tilde{a}_0 = a_0 \cdot (1-\rho)$; and

$\tilde{X}_{i,t-1} = X_{i,t-1} - \rho \cdot X_{i,t-2}$

and the predicted model:

$$\hat{\tilde{Y}}_{t+1} = \tilde{a}_o + \hat{a_1} \cdot \tilde{X}_{1,t} + \hat{a}_2 \cdot \tilde{X}_{2,t} + \hat{a}_3 \cdot \tilde{X}_{3,t}$$

then,

$$\hat{Y}_{t+1} = \hat{\tilde{Y}}_{t+1} + \rho \cdot Y_t. \tag{5.14}$$

The question is how to estimate ρ? One method suggested by Durbin (in Maddala [13]) is to regress,

$$\begin{aligned} Y_t &= a_o(1-\rho) + \rho \cdot Y_{t-1} + a_1 \cdot X_{1,t-1} - a_1 \cdot \rho \cdot X_{1,t-2} \\ &\quad + a_2 \cdot X_{2,t-1} - a_2 \cdot \rho \cdot X_{2,t-2} + a_3 \cdot X_{3,t-1} \\ &\quad - a_3 \cdot \rho \cdot X_{3,t-2} + e_t \end{aligned} \tag{5.15}$$

and use $\hat{\rho}$ as the estimate of ρ.

Therefore, by using AR(1) model and substituting Y_t's and $X_{i,t}'$s for the variables in this study, the following model is obtained:

$$LNMP87 = a_o' + \gamma \cdot LNMP76 + a_1' \cdot LNPp70 - a_1' \cdot \gamma \cdot LNPp60 \tag{5.16}$$

$$+a_2' \cdot LNUp70 - a_2' \cdot \gamma \cdot LNUp60 + a_3' \cdot LNIp70$$
$$-a_3' \cdot \gamma \cdot LNIp60 + e_t$$

Applying the AR(1) model above into this study, regress,

$$\begin{aligned} RLNMP87 &= a_o + a_1 \cdot RLNPp76 + a_2 \cdot RLNUp76 \\ &\quad + a_3 \cdot RLNIp76 + e_t. \end{aligned} \tag{5.17}$$

where, $RLNMP87 = LNMP87 - \hat{\gamma} \cdot LNMP76$;

$RLNPp76 = LNPp70 - \hat{\gamma} \cdot LNPp60$;

$RLNUp76 = LNUp70 - \hat{\gamma} \cdot LNUp60$;

$RLNIp76 = LNIp70 - \hat{\gamma} \cdot LNIp60$; and

$\hat{\gamma}$ = the estimate of γ, i.e., the coefficient of LNMP76 in the regression 5.16.

From the above discussions, those six initial models will be compared and the results are discussed in the following chapter.

6. RESULTS, EFFICIENT MODEL AND PREDICTION

6.1 Initial Results

By applying different assumptions discussed in the previous chapter, some initial results can be obtained as shown in Table 6.1.

Among the seven models, no model has all coefficients of regressors significantly different from zero at 5 % significance level testing, and the $\bar{R}^2$ is quite low. Models like BASIC, RL, AR(1), SUBT and DIFF have $\bar{R}^2$ not exceed 0.50. This may be because the low explanatory ability to per capita in-migration (M(70)/P(60)) on the regressors Pp(60), Ip(60) and Up(60). The $\bar{R}^2$ of that in RL model has only 0.1843. That means over 80 % of the variance of the dependent variable has not been explained. This also provides a chance for the variable RES76 (the residual by regressing M(70)/P(60) on the regressors LNPp60, LNIp60 and LNUp60) in the RES model to make a big contribution to the M(80)/P(70) regression. Since a lot information remains unexplained, within the residual, RES76 provides some non-random explanation to the variation of the M(80)/P(70) observations. Although the sources of variation can be detected to be related to the two regressions, the explanatory abilities of those variables which significant but missing from the RL model are retained in the RES76. This can be seen by the high $\bar{R}^2$ of 0.8557 in RES model. The disadvantages of this model is that, the RES76 by definition is a random error of the RL model, if the RL model is true, then RES76 will contain only random disturbances. Although what the RES76 may now contain not just the random error, but also something which is excluded from the RL model. In short, the RES76 in the RES model still contains some random effects which are not preferred in any explanatory variable. On the other hand, by employing RES76, the actual source of variations cannot be determined.

For the RLL model, it is free from the bad performance of the past regression on the 70's data because it takes the past dependent variables straight into the model. Since RES76 = LNMP76 - $LN\widehat{MP}76$, high $\bar{R}^2$ obtained in both cases are expected. In the RLL model, $\bar{R}^2$ equals to 0.8557, but the regressors LNUp70 and LNIp70 are insignificantly different from zero in the test of 5 % significance level. Someone might raise a question to the nature of autoregressive process in the model produce inconsistent estimates, the variance might not be stationary over time. But note that this is not a time series study. In fact, this is regression analysis by using the information of the past three records to predict the behavior of a period ahead. Time series technique is not applicable. Thus, we can see the LNMP76 as an independent variable because it is not related to any lagged term in the past. That is, the regression models of LNMP76 and LNMP65 are different models with different coefficients from the model discussed in this study.

The last thing to do in this procedure is to choose between the RES and the RLL to be the initial model. Although RES model resulted in a slightly higher $\bar{R}^2$ by 0.0032 than that of the RLL, the RLL model is still chosen to be the best initial model. The main reason is the RES76 contains random effects while the lagged dependent variable in the RLL model is a fixed, independent and observable variable. Those are preferred conditions in the Ordinary Least Square Estimation. Another reason is that the results of the RLL model are much better and easier to be interpreted.

6.2 Refining the initial model

From the RLL model, the useful information for using the regressor LNMP76 is obtained, the next question is how can the model be improved?

First of all, seven simple regressions are obtained by regressing the LNMP87 on each of the possible regressors. The results are listed on Table 6.2.

Then, according to the performance of explanatory ability, the regressors are added into the model one by one. By reviewing the increment of $\bar{R}^2$ and the significance of the regressor, the best model subject to the restricted economic variables is obtained. The best model is determined based on the model's performance on:

(i) Increment on $\bar{R}^2$;

(ii) Significance of regressors at 5% significance level testing;

(iii)Mallow's Cp[1] statistics of overall model; and

(iv) Condition number bounds:

CN_u = Condition Number upper bound statistic;

CN_l = Condition Number lower bound.

For the application and criteria of Cp statistic and condition number bounds, see Rawlings [16].

A summary of model refining is listed on Table 6.3. It shows the models which are the highest $\bar{R}^2$ models compared to the those have the same number of regressors.

[1]Mallow suggested models with Cp statistics close to the number of coefficients can be the candidates for the efficient model.

The model RLL1 inputed only the lagged dependent variable, LNMP76, as the explanatory variable, this model can be so called the naïve model. Naïve hypotheses some dependent variable is best explained by its lagged term. It shows that, in this study, the LNMP76 accounted for 80% of the total variations of the LNMP87. However, the Cp statistics suggests that this may not be the best model. Therefore, LNPp60 is added to RLL1 to form model RLL2. $\bar{R}^2$ increased to 0.8554 in RLL2.

When the regressors are continued to be added into the models, the Cp statistics falls from 154.1 to a minima of 4.0. On the other hand, the $\bar{R}^2$ increases from 0.8012 up to maxima of 0.8610 at the model RLL7. Can it determine model RLL7 is the best model? Of course not. Although Cp minimized and $\bar{R}^2$ maximized at RLL7, CN_l of 304.7 indicates serious multi-collinearity among the regressors. In addition, LNUp60 is insignificantly different from zero, so RLL7 cannot be the best model. Thus, the next choice will be fallen to RLL5 and RLL6 because they both have almost the same $\bar{R}^2$, Cp statistics and condition number bounds and, more important, the coefficients are all significant in RLL5. Consider the signs of the coefficients obtained, they are all in the expected signs. LNMP76, LNIp70, and LNUp60 are of "+" signs while LNPp60, LNPp70, LNUp70 and LNIp60 have "-" signs.

Consider the RLL7 and RLL8, LNIp60 and LNIp70 are significant in both models but have opposite signs. This suggests that the differential of log income potential may be an improvement to the model. Therefore, LNIpD76 (=LNIp70 - LNIp60) is introduced to the initial RLL model and replaced LNIp70 and LNIp60. Redo the model refining procedure, models which attain highest $\bar{R}^2$ with different number of regressors inputed are shown on Table 6.4.

From the second refining stage, RLL12, RLL13 and RLL14 contain insignificant variables, so they cannot be the best model. RLL11 can be chosen to be the best model because,

(i) its highest $\bar{R}^2 = 0.8603$ among all possible models (RLL9, RLL10 and RLL11);

(ii) although its Cp statistics = 4.0 > 3 = its number of regressors, but it has a very sharp fall among the feasible models;

(iii) its CN_l and CN_u both are small enough to state that multi-collinearity among regressors is not significant;

(iv) all coefficients are significantly different from zero; it indicates that each of them can explain somehow the variation of the dependent variable;

(v) all coefficients are of expected signs;

(vi) $F = 742.02 > F_{(3,358),.95} = 2.60$ indicates the whole regression is significant at 5 % significance level; and

(vii) the normal plot of residual, Figure 6.1, looks straight, it indicates the residuals are normally distributed.

Therefore, the best model (RLL11) is

$$\begin{aligned} LNMP87_i &= 2.98982848 + 0.76417338 \cdot LNMP76_i \\ &\quad -0.30016646 \cdot LNPp70_i \\ &\quad +1.84658321 \cdot LNIpD76_i + V_i \end{aligned} \tag{6.1}$$

where, $i = 1, 2, \cdots, 362$; $V_i \sim i.i.d.(0, \sigma_v^2)$.

Although this model has $\bar{R}^2$ at about 0.85 which is about 5 % less than Greenwood's or Mueser's model, one should notice that this model contains only three regressors and all of them can be obtained directly from the past records. This model is much more simple than Greenwood's simultaneous equation system.

6.3 Interpretation of the model

From the model RLL11, there are three important findings. First of all, the Up shows insignificant effect in the in-migration decision. One reason to explain this phenomenon is that the unemployment rate differentials across all CRs are not serious enough to stimulate in-migration flow. Even being unemployed labors may not migrate to a new region if that region cannot provide much better chance for him to find a job. This may not due to the unemployment potential made the unemployment rate more evenly distributed across the whole country. Remember that the duration for the new came labor to wait in the unemployment pool for a job is also one of the costs that he needs to consider before taking the move. Recall the correlation matrix in Table 3.1, someone might suspect that Up is highly correlated with Pp and Ip, therefore, some of the Up's effect might have been explained by them. It probably is not in this case. It is because, although U, the unemployment rate without taking potential transformation, are not seriously correlated with P and I, but this phenomenon can still be observed on the alternative models, see Table 6.5 in the next section. Unemployment rate can never significant in the best model in any case.

Another reason might provide an explanation about the significance of Up is that the nature of the unemployment rate is just an index to show the percentage of people being unemployed, it usually cannot tell how many carpenters or accountants are being unemployed in this generalized figure. It cannot distinguish which markets are saturated or which are not. In fact, the labors in the real world are not identical and the information is not perfectly known, the assignment problems cannot be solved easily. The observed unemployment rate in the U.S. could not stimulate in-migration during the studied period.

The second finding in this study is that the in-migration depends on the growth of income rather than the relative income. Note that the coefficient of the variable LNIpD76 is 1.8466. Recall that the LNIpD76=LNIp70 - LNIp60 (or $\ln \frac{Ip70}{Ip60}$), it is just the natural log of the growth rate of income potential from 1960's to 1970's data. That means the migrants are not looking for short run high income to determine their destinations, instead, they consider the growth of it. This is quite reasonable because the higher the growth of income in one place, it means the income increased during the past ten years in that place is higher than in other places. This is quite a longer term consideration, and the migrants look for higher income in the future. Therefore, the higher the growth of income induced stronger attractional force to in-migration is quite sure.

Thirdly, this study suggests the use of the lagged dependent variable as the regressor. This implies the future per capita in-migration flow is partially induced by the present per capita in-migration, or simply, future in-migration is partially induced by the present in-migration. It is also true because most migrants are not likely to move to a completely new place, they usually like to obtain information and experience from the relatives or friends who live there before making decision. This channel of information flow is so efficient that future in-migration can be expected to follow certain trend of the present in-migration. Of course, it can happen only if a steady economic environment, and no other big unexpected changes in the U.S.

The people continue moving away from the highly populated areas can be observed in the 1980s. The coefficient of Pp signed negative. The higher the Pp, the less the in-migration will be.

Recall that the structural form of the RLL model is a Cobb-Douglas Equation, that means the coefficient of each of the variables by definition is the elasticity, ζ, of that variable to the per capita in-migration (MP87). That is,

$$\zeta_{MP} = \frac{\partial LNMP87}{\partial LNMP76} = \frac{\%\Delta MP87}{\%\Delta MP76} \doteq 0.7642$$

$$\zeta_{Pp} = \frac{\partial LNMP87}{\partial LNPp70} = \frac{\%\Delta MP87}{\%\Delta Pp70} \doteq -0.3002$$

$$\zeta_{IpD} = \frac{\partial LNMP87}{\partial LNIpD76} = \frac{\%\Delta MP87}{\%\Delta IpD76} \doteq 1.8466$$

That means a 1 % increase (decrease) in the MP76, keeping other things unchanged, there will be an approximate of 0.7642 % increase (decrease) in the MP87. Same implication can be applied to the variable LNIpD76 that 1 % increase (decrease) in the grow of income potential will induce 1.8466 % increase (decrease) in MP87. In the case of LNPp70, the ζ_{Pp} implies a 1 % increase (decrease) in the Pp70, keeping other things unchanged, there will be an approximate of 0.3002 % decrease (increase) in the MP87.

6.4 The Need of Potentials

To compare the need of potentials, four alternative initial RLL models are studied by using the observed economic variables (P, I and U). That is,

(1) Regress LNMP87 on P70, P60, I70, I60, U70, U60 and LNMP76;

(2) Regress MP87 on P70, P60, I70, I60, U70, U60 and MP76;

(3) Regress LNMP87 on LNP70, LNP60, LNI70, LNI60, LNU70, LNU60 and

LNMP76; and

(4) Regress MP87 on LNP70, LNP60, LNI70, LNI60, LNU70, LNU60 and MP76; where, MP87 = M80/P70;

P70 = Population of 1970, etc; and

LNP70 = Natural log of P70, etc.

A summary of the best models of the four alternative initial RLL models are listed on Table 6.5.

From the four regressions, the following facts are observed:

Fact 1: Best models have MP87 as dependent variable can only maximize the $\bar{R}^2$ at 0.75 which is 0.1 inferior to that of the RLL11.

Fact 2: Best models have LNMP87 as dependent variable cannot obtain expected sign on the coefficient of I or LNI. The negative coefficient suggest that the people move to the lower per capita real income regions. That cannot reflect the behavior of the real world.

Having the above evidence, the usefulness and the effectiveness of potential analysis on in-migration can be verified. The use of potentials as regressors can, on the one hand, obtain higher explanatory ability; on the other hand, it keeps the logical relationship between regressors and the dependent variable.

6.5 Prediction

The prediction equation of the model RLL11 of the year 1990 is:

$$\begin{aligned} LN\widehat{MP}98_i &= 2.98982824 + 0.76417338 \cdot LNMP87_i \\ &\quad -0.30016646 \cdot LNPp80_i \end{aligned} \tag{6.2}$$

$$+1.846588321 \cdot LNIpD87_i.$$

Since,

$$LN\widehat{M}P98_i = \ln(\frac{\widehat{M90}_i}{P80_i})$$

so, the predicted in-migration of CR_i in 1990 ($\widehat{M90}_i$) will be

$$\widehat{M90}_i = P80_i \cdot \exp(LN\widehat{M}P98_i).$$

The estimated $\widehat{M90}_i$, $\widehat{M80}_i$ and the observed $M80_i$ are plotted together on Figure 6.2 and the estimated values are listed in the Appendix F.

Table 6.1: Regression Results of Initial Models

	Initial Models								
	BASIC		RL		RLL	SUBT	RES	AR(1)	DIFF
	Dependent Variable								
LNMP87	⋆		⋆		⋆	⋆	⋆		⋆
LNMP76		⋆		⋆					
RLNMP87								⋆	
	Regressors								
LNPp80	⋆								
LNPp70		⋆	⋆		⋆	⋆	⋆		⋆
LNPp60				⋆		⋆			
LNUp80	⋆								
LNUp70		⋆†	⋆†		⋆†	⋆†	⋆†		⋆†
LNUp60				⋆		⋆			
LNIp80	⋆								
LNIp70		⋆†	⋆†		⋆†	⋆	⋆		⋆†
LNIp60				⋆		⋆			
LNMP76					⋆				
RES76							⋆		
RLNPp76								⋆	
RLNUp76								⋆†	
RLNIp76								⋆†	
LNMPD87									⋆†
R^2	.3450	.1368	.3047	.1775	.8557	.4531	.8557	.3006	.3045

⋆: Variable which is used in the regression. † : Variable whose coefficient is insignificantly different from zero in a 5 % significance level t test.

Table 6.2: Simple Regressions of LNMP87 on Regressors

REGRESSOR	R^2
LNMP76	0.8017
LNPp70	0.3013
LNUp70	0.2139
LNIp70	0.2430
LNPp60	0.3174
LNUp60	0.2602
LNIp60	0.2560

Figure 6.1: Normality Plot of the Residuals of RLL11

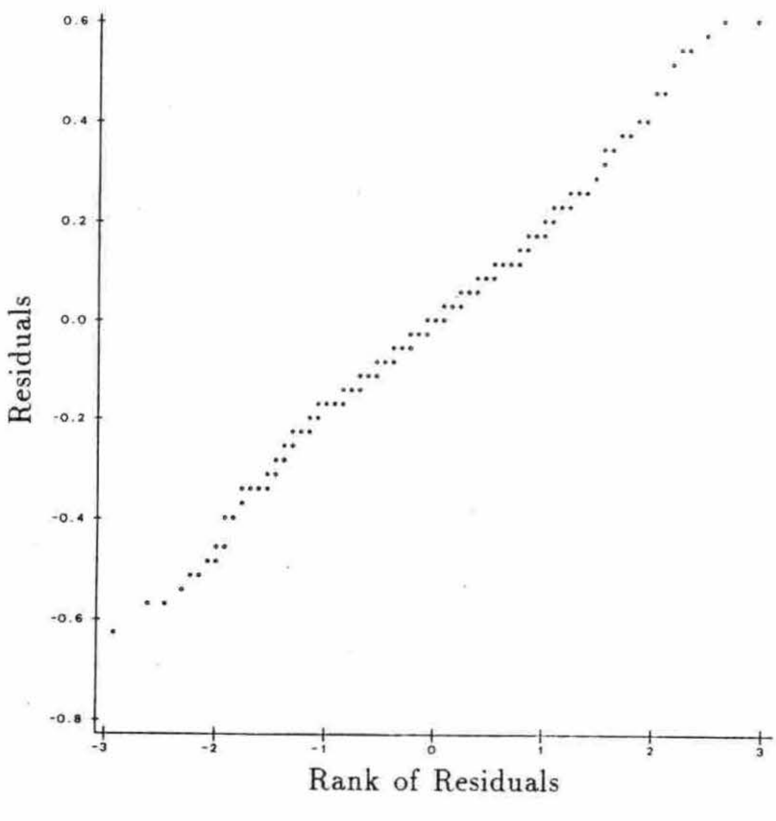


Table 6.3: Summary of Model Refining (Stage 1)

	MODEL							
VARIABLE	RLL1	RLL2	RLL3	RLL4	RLL5	RLL6	RLL7	RLL8
LNMP76	.8658	.7756	.7776	.7863	.7790	.7809	.7709	.7707
	(38.2)	(37.2)	(37.5)	(37.0)	(36.6)	(36.4)	(36.5)	(34.9)
LNPp70		-.3070	-.4621	-.4829			-.3732	-.3723
		(-12.)	(-6.5)	(-6.7)			(-4.6)	(-4.1)
LNUp70				$-.3311^\dagger$	-.3884	$-.3066^\dagger$		$.0057^\dagger$
				(-1.8)	(-2.1)	(1.46)		(.03)
LNIp70							1.5412	1.5474
							(2.9)	(2.7)
LNPp60					-.4907	-.4418		
					(-6.7)	(-4.8)		
LNUp60			.2321	.5592	.6382	.6910	$.3185^\dagger$	$.3141^\dagger$
			(2.34)	(2.69)	(2.98)	(3.1)	(1.76)	(1.25)
LNIp60						$-.1752^\dagger$	-1.726	-1.734
						(-.86)	(-3.4)	(-2.9)
INTERCEPT	-.1045	3.6238	4.2769	4.5281	4.4216	4.4079	3.6277	3.6210
$\bar{R}^2$	.8012	.8554	.8572	.8581	.8583	.8582	.8610	.8606
Cp	154.1	15.3	11.6	10.4	9.9	11.2	4.0	6.0
CN_u	1.0	4.6	54.9	289	304	683	3188	5196
CN_l	1.0	1.2	8.6	37.6	40	45	305	394

The numbers in "( )" are computed t values, for 5 % significance level, the critical t boundaries are (-1.96, +1.96).

†: the coefficient of regressor is insignificantly different from zero in the t test.

Table 6.4: Summary of Model Refining (Stage 2)

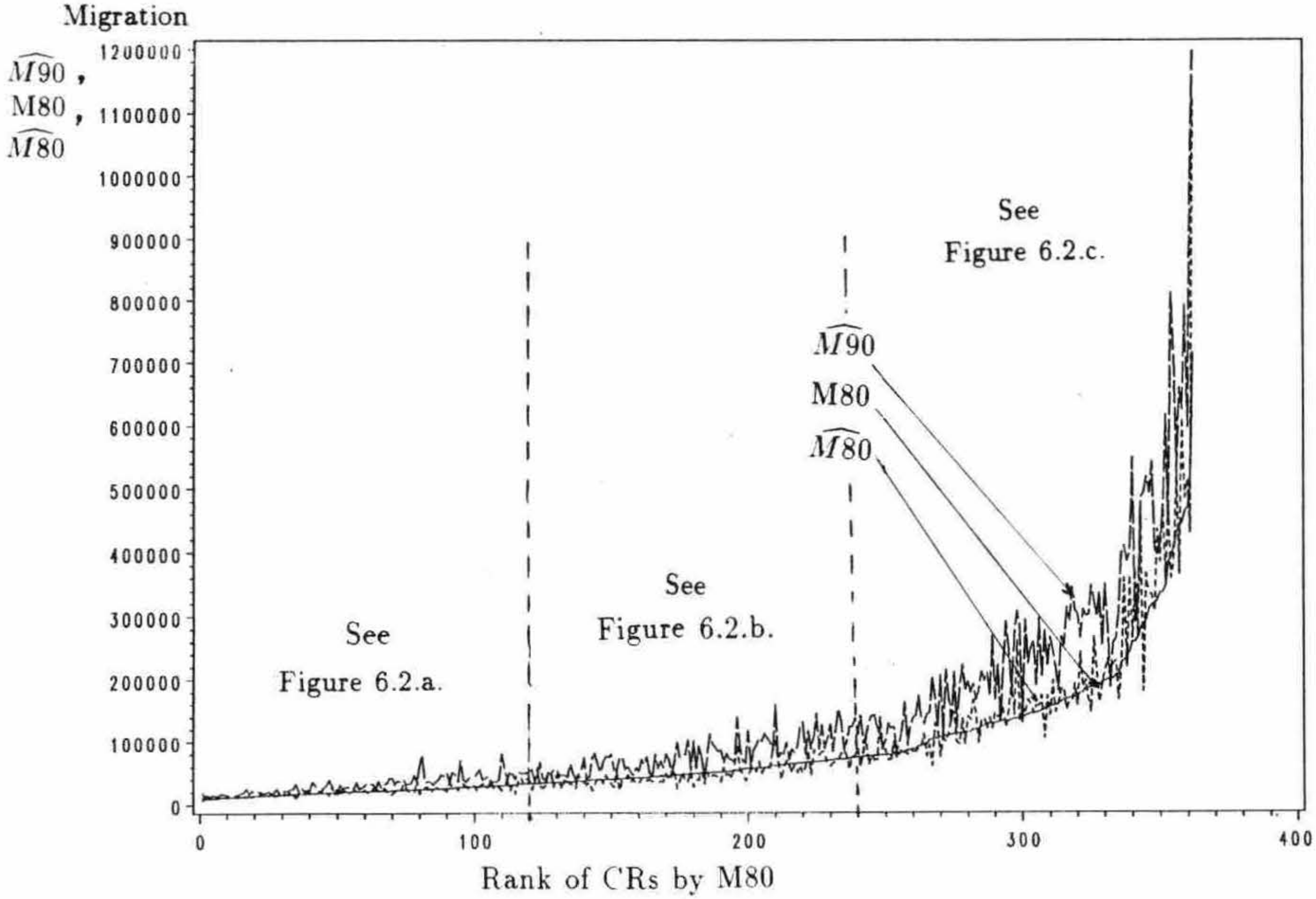
VARIABLE	RLL9	RLL10	RLL11	RLL12	RLL13	RLL14
LNMP76	.8658	.7756	.7642	.7666	.7632	.7661
	(38.1)	(36.6)	(36.8)	(37.0)	(34.8)	(32.8)
LNPp70		-.3070	-.3002	-.4124		−.209†
		(-11.6)	(-11.6)	(5.74)		(-.36)
LNUp70					.1237†	−.0971†
					(-.59)	(-.44)
LNPp60					-.4157	−.215†
					(-5.6)	(-.37)
LNUp60				.1670†	.3110†	.2779†
				(1.67)	(1.27)	(1.06)
LNIpD76			1.8466	1.6761	1.5433	1.5614
			(3.67)	(3.27)	(2.65)	(2.67)
INTERCEPT	-.105	3.624	2.990	3.518	3.531	3.581
$\bar{R}^2$	.8012	.8551	.8603	.8610	.8606	.8603
Cp	154.4	15.5	4.0	3.2	5.1	7.0
CN_u	1.0	4.6	10.2	80.7	497	7885
CN_l	1.0	1.2	1.2	9	53.6	635

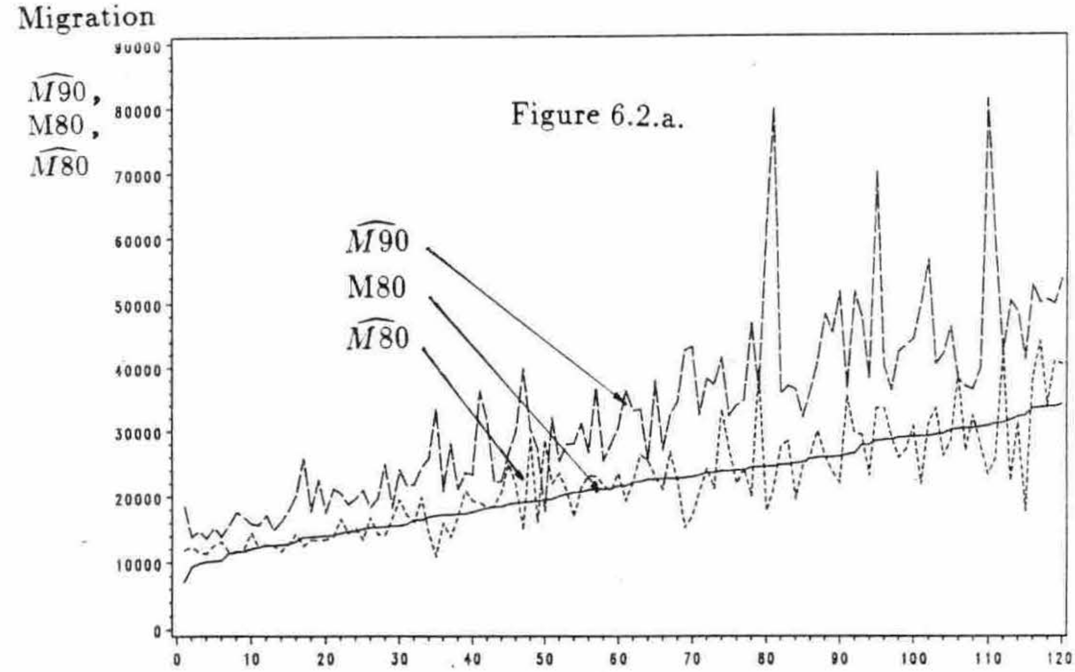
See footnote on Model Refining (Stage 1).

Table 6.5: Summary of the Alternative Initial Models

	Dependent Variable			Dependent Variable	
Regressor	LNMP87	MP87	Regressor	LNMP87	MP87
LNMP76	.8682		LNMP76	.8869	
MP76		.8331	MP76		.8704
LNP70	-.0635	-.0191	P70	-6E-8	-1E-8
LNP60			P60		
LNI70	-.3980	-.0517	I70	-8E-5	-1.1E-5
LNI60			I60		
LNU70			U70		
LNU60			U60		
Constant	4.0891	.7372	Constant	.3737	.1154
$\overline{R}^2$	.8451	.7631	$\overline{R}^2$	.8408	.7505
Cp	4.2	1.8	Cp	1.8	1.0
CN_u	11.7	10.7	CN_u	11.5	10.6
CN_l	1.4	1.2	CN_l	1.3	1.2

Figure 6.2: Predictions of In-migration of 1980 and 1990





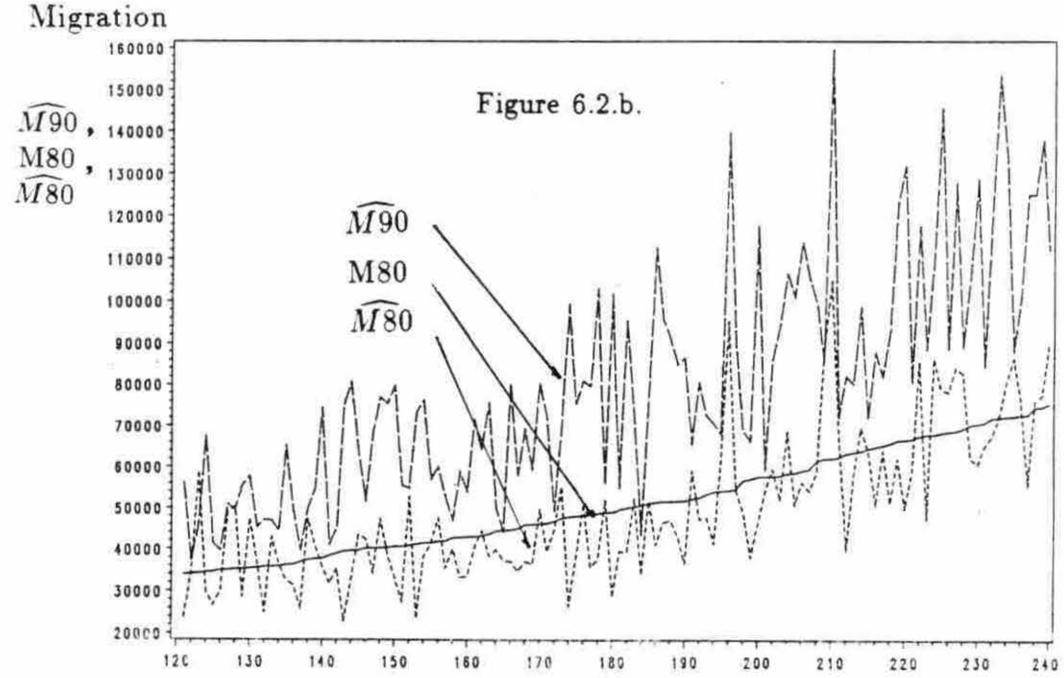


Figure 6.2 (Continued)

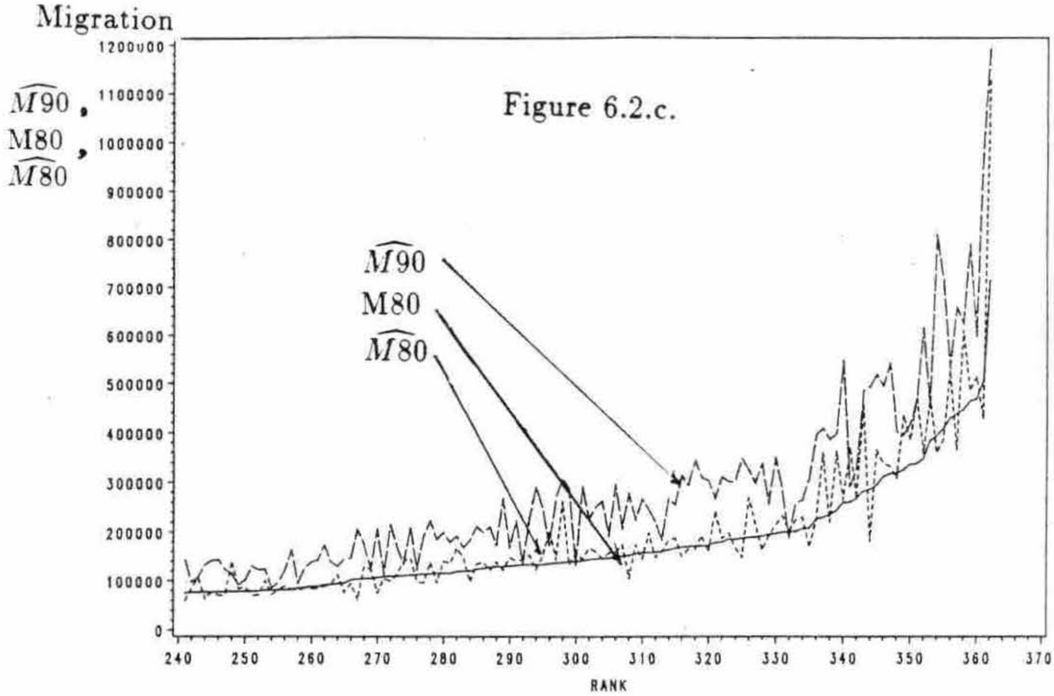


Figure 6.2 (Continued)

7. LIMITATIONS AND FURTHER APPLICATIONS

When doing this research, there were many restrictions on the data structure, on the assumptions of variables and on the prediction of the growth of new cities, which limited the predictability and validity of the model chosen.

7.1 The Data Structure

The dataset is collected according to the fixed CRs, each observation has to be regrouped from the county's data from the County and City Data Books and the Census Reports. In the 1950 census, the median income was used instead of per capita income and the residence shifts definition changed from five years to one year (after 1950, the definition was switched back to five years). This restriction stopped the data collection at 1960. Therefore, it limits the use of time series analysis method from this study. This restricted the analysis within a time span of last thirty years.

7.2 The Assumptions

To calculate the per capita real income, the real income has to be deflated from the current income data. Since price levels varied not only among different regions but also among different cities within the same region, the regions across city sizes consumer price indices were found from the consumer price index detailed reports.

The CPI(U) was newly introduced on Jan., 1978 and so no data on this item before that year. It is necessary to assume that the price level pattern of regions across city sizes are the same during 1960 to 1980. In fact, by comparing the reported figures, price indices were quite stable in values and relatively across city size. So the assumption made above is not unreasonable. However, some abnormal cases which the price level fluctuated seriously year to year are excluded.

For the Regional Price Index, some assumptions are made such as the fixed consumption pattern of the people, and also due to the data source limitation, the CRs of size "C" and the rural areas (size "D") are assumed to have two indices regardless the region.

On the other hand, distances among CRs are assumed to be constant over time. To obtain the perfect potential measures, it is ideal to release the constant distance assumption and obtain the distance changes at 60s, 70s and 80s. Comparing the 1973 edition and the 1985 edition of the S.H.M.G., the cities appeared on both editions have only small mileage changes, such as 2 miles changes over 1080 miles distance apart between Bloomington at Indiana and Boulder at Colorado. These may be caused by the growth of the cities and so the shift of the city centers influenced the measures. In addition, problems arise only on the new grown cities, which have highways newly built to link with other cities. Another question is how to trace back the transportation distance between two places before the highway is built? To solve the this question, it is not impossible, but it might have errors in deciding the center's location, and the routes to be used in the measures. So, it is better to ask how many highways were built during the period of study, if not a large number, it might be benefit to use an assumption of constant distance to replace those uncertain measures.

7.3 Boundary Definition

This study fixed the CRs by the definitions of MCRs and NCRs, they altogether occupied every inch of the U.S. One might ask about the new grown cities; may be 20, may be more new grown cities are detected at each census. Can this study provide any hint for the locations of the new city? A new growth city existence depends on the change of total population in the area. The change of population is consisted of the number of birth changes, the number of death changes and the net-migration changes. If it is assumed that the birth and death rates are almost constants in a closed society, then this model can tell half of the story about the development of new cities. The other half depends on the out-migration. Therefore, if the dependent variable of this study changes to the volume of net migration (net change of migration), then the predicted net migration volume can approximately tell the possible CR where new cities will be developed, the same structural model can be applied by using micro-data obtained within the CR and subdivided the CR into smaller units, such as counties, and sub-districts. Nevertheless, those are the extended studies suggest to the state or regional governments that will not describe in detail here.

Another point to be noticed by fixing the CR boundaries is on the commuters across the CR boundaries. There is no restriction to the residents of CR_a cannot enter CR_b, and vise versa. So, the in-migration change of CR_a suggests not only a change of the infra-structures and public facilities for CR_a is required, but also that for the neighboring CRs. Therefore, the state or regional governments might also study the neighboring CRs before making budgeting decisions. In fact, the in-migration of one CR change may be due to the economic or environmental changes of its neighboring CRs. Thus, in-migration problem is a dynamic study and cannot be studied statically.

The purpose of this study is only to suggest an estimated pattern of in-migration shifting in the years 1985 to 1990. For the in-migration shifting pattern in the years 1995 to 2000, readers might want to re-do the same study by providing either a new CR definition or the one used here, but with the observed datasets of the year 1990.

7.4 Exceptional Cases

Since this study based on the economic environment view point to explain in-migration behavior, there are many non-economic external factors which influence in-migration decision, such as, personal reasons, social reasons, political reasons and geographical reasons. For example, some in-migration decisions may be changed after the earthquake in San Francisco in the West, or after the hurricane attacked in the North Carolina in the East early this year, may be, some people had been planned to move towards the central part of the continent. All such geographic factors triggering psychological effects on in-migration actions are hard to be predicted and are out of the boundaries of economic reasons. Therefore, those cases are not included in this study.

7.5 Notes

The potential measures can be a measure of potential by using totals: total regional real income and total regional unemployment. Together with the population potential, those variables can be used to form the similar analysis. Actually, I had tried to use those potential of totals redo the analysis, the results are also the same and are briefly described as follows: (1) the coefficients are significant but in very small values; (2) the total unemployment potential was inserted instead of the income growth potential in the best model; and (3) the adjusted R^2 of the best models are almost identical for the two analyses.

8. CONCLUSION

This study is aimed to introduce potential analysis into econometric modeling and to find an effective model to represent and predict regional migration flow. The findings show that the application of potentials is essential to form an efficient model in this study.

The advantages of the best model, RLL11, are (1) it is consisted of only three variables; (2) its regression result is superior to that of not using potential form regressors in the regression; and (3) it requires only the variables in the time t and t-10 to make prediction for the in-migration flow in the time t+10. As one of the discovery in this study, unemployment rate of the control regions is insignificant to influence in-migration decisions. Another findings is that the past in-migration and the income differential are important variables to explain in-migration behavior.

This study has experimented to insert spatial relationship into regressors to form an effective and efficient spatial interaction model. The internal in-migration regression done here verified the potential analysis hypothesis is applicable. Potential analysis can be broadly used to spatial analysis in regional studies.

9. BIBLIOGRAPHY

[1] T. R. Anderson. *Intermetropolitan Migration: A Comparison of the Hypaotheses of Zipf and Stouffer.* American Sociological Review, 20, June, 1955, 287-291.

[2] T. R. Anderson. *Potential Models and the SpatialDistribution of Population.* Paper and Proceedings of the Regional Science Association, 2, 1956, 175-182.

[3] G. A. P. Carrothers. *An Historical Review of the Gravity and Potential Concepts of Human Interaction.* Journal of American Institute of Planners, 22, Spring 1956, 94-102.

[4] G. A. P. Carrothers. *Population Projection by Means of Income Potential Models.* Paper and Proceedings of the Regional Science Association, 4, 1958, 121-160.

[5] E. S. Dunn. *The Market Potential Concept and the Analysis of Location.* Paper and Proceedings of the Regional Science Association, 2, 1956, 183-198.

[6] D. K. Foot and W. J. Milne. *Multiregional Estimation of Gross Internal Migration Flows.* Internation Regional Science Review, 12, No. 1, 1989, 29-43.

[7] M. J. Greenwood. *A Simultaneous-Equations Model of Urban Growth and Migration.* Journal of the American Statistical Association, Vol. 70, No. 352, Dec. 1975, 797-810.

[8] M. J. Greenwood and G. L. Hunt. *Migration and Interregional Employment Redistribution in the United States.* The American Economic Review, Vol. 74, No. 5, Dec. 1984, 957-969.

[9] W. Isard. *Method of Regional Analysis.* M.I.T. Press, 1960.

[10] G. G. Judge, R. C. Hill, W. Griffiths, H. Lütkepohl and T. C. Lee. *Introduction to the Theory and Practice of Econometrics.* John Wiley and Sons, 1982.

[11] E. H. Jurkat. *Discussion: Geography of Price and Spatial Interaction.* Paper and Proceedings of the Regional Science Association, 3, 1957, 130-133.

[12] J. B. Kau and C. F. Sirmans. *A Recursive Model of the Spatial Allocation of Migration.* Journal of Regional Science, Vol. 19, No. 1, 1979, 47-56.

[13] G. S. Maddala. *Introduction to Econometrics.* Macmillan Publish, 1988.

[14] P. Mueser. *The Spatial Structure of Migration: An Analysis of Flows Between States in the USA over Three Decades.* Regional Studies, Vol. 23.3, June 1989, 185-200.

[15] Rand McNally and Company. *Standard Highway Mileage Guide.* Chicago: Rand McNally and Company, 1985.

[16] J. O. Rawlings. *Applied Regression Analysis: A research tool.* The Wadsworth and Brooks, 1988.

[17] A. Schwartz. *Interpreting the Effect of Distance on Migration.* Journal of Political Economy, Vol. 81, No. 5, 1973, 1153-1169.

[18] J. Q. Stewart. *Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population.* Geographical Review, 37:3, July 1947, 461-485.

[19] J. Q. Stewart. *Potential of Population and Its Relationship to Marketing.* in Theory of Marketing, Chicago: Richard D. Irwin, 1950, 19-40.

[20] J. Q. Stewart and W. Warntz. *Physics of Population Distribution.* Journal of Regional Science, 1, Summer 1958, 99-123.

[21] U. S. Bureau of Census. *County and City Data Book.* Washington. D. C.: U. S. Government Printing Office, 1965, 1972, 1986.

[22] U. S. Bureau of Census. *Census of Population: Characteristics of Population.* Washington. D. C.: U. S. Government Printing Office, 1960, 1970, 1980.

[23] A. M. Voorhees. Discussion: Geography of Price and Spatial Interaction. Paper and Proceedings of the Regional Science Association, 3, 1957, 130-133.

[24] W. Warntz. *Geography of Prices and Spatial Interaction.* Paper and Proceedings of the Regional Science Association, 3, 1957, 118-129.

[25] W. Warntz. *Toward a Geography of Price.* Universitu of Pennsylvania Press, 1959.

10. APPENDIX A: LIST OF CRs

Table 10.1: List of CRs

Code	Name	State	Code	Name	State
1	Abilene	TX	24	Bakersfield	CA
2	Akron	OH	25	Baltimore	MD
3	Albany	GA	26	Bangor	ME
4	Albany-	NY	27	Banton Rouge	LA
	Schenectady-		28	Battle Creek	MI
	Troy		29	Beaumont-	TX
5	Albuquerque	NM		Port Arthur	
6	Alexandria	LA	30	Beaver County	PA
7	Allentown-	PA-NJ	31	Bellingham	WA
	Bethlehem		32	Benton Harbor	MI
8	Altoona	PA	33	Bergen-	NJ
9	Amarillo	TX		Passaic	
10	Anaheim-	CA	34	Billings	MT
	Santa Ana		35	Biloxi-	MS
11	Anchorage	AK		Gulfport	
12	Anderson	IN	36	Binghamton	NY
13	Anderson	SC	37	Birmingham	AL
14	Ann Arbor	MI	38	Bismarck	ND
15	Anniston	Al	39	Bloomington	IN
16	Appleton	WI	40	Bloomington-	IL
	Oshkosh-			Normal	
	Neenah		41	Boise City	ID
17	Asheville	NC	42	Boston-	MA
18	Athens	GA		Lawrence-	
19	Atlanta	GA		Salem-	
20	Atlantic City	NJ		Lowell-	
21	Augusta	GA-SC		Brockton	
22	Aurora-	IL	43	Boulder-	CO
	Elgin			Longmont	
23	Austin	TX	44	Bradenton	FL

Table 10.1 (Continued)

Code	Name	State	Code	Name	State
45	Brazoria	TX	65	Clarksville-	TN-KY
46	Bremerton	WA		Hopkinsville	
47	Bridgeport-	CT	66	Cleverland	OH
	Milford-		67	Colorado Sps.	CO
	Stamford-		68	Columbia	MO
	Norwalk-		69	Columbia	SC
	Danbury		70	Columbus	GA-AL
48	Brownsville-	TX	71	Columbus	OH
	Harlingen		72	Corpus Christi	TX
49	Bryan-	TX	73	Cumberland	MD-WV
	College Station		74	Dallas	TX
50	Buffalo	NY	75	Danville	VA
51	Burlington	NC	76	Davenport-	IA-IL
52	Burlington	VT		Rock Island-	
53	Canton	OH		Moline	
54	Casper	WY	77	Dayton-	OH
55	Cedar Rapids	IA		Springfield	
56	Champaign-	IL	78	Daytona Beach	FL
	Urbana-		79	Decatur	IL
	Rantoul		80	Denver	CO
57	Charleston	SC	81	Des Moines	IA
58	Charleston	WV	82	Detroit	MI
59	Charlotte-	NC-SC	83	Dothan	AL
	Gastonia-		84	Dubuque	IA
	Rock Hill		85	Duluth	MN-WI
60	Charlottesville	VA	86	Eau Claire	WI
61	Chattanooga	TN-GA	87	El Paso	TX
62	Chicago	IL	88	Elkhart-	IN
63	Chico	CA		Goshen	
64	Cincinnati	OH-KY-IN	89	Elmire	NY

Table 10.1 (Continued)

90	Elrid	OK		Texas City	
91	Erie	PA	112	Gary-	IN
92	Eugene-	OR		Hammond	
	Springfield		113	Glens Falls	NY
93	Evansville	IN-KY	114	Grand Forks	ND
94	Fargo-	ND-MN	115	Grand Rapids	MI
	Moorhead		116	Great Falls	MT
95	Fayetteville	NC	117	Greeley	CO
96	Fayetteville-	AR	118	Green Bay	WI
	Springdale		119	Greensboro-	NC
97	Flint	MI		Winston-	
98	Florence	AL		Salem-	
99	Florence	SC		High Point	
100	Fort Collins-	CO	120	Greenville-	SC
	Loveland			Spartanbury	
101	Fort Lauderdale-	FL	121	Hagerstown	MD
	Hollywood-		122	Hamilton-	OH
	Pompano Beach			Middletown	
102	Fort Myers-	FL	123	Harrisburg-	PA
	Cape Coral			Lebanon-	
103	Fort Pierce	FL		Carlisle	
104	Fort Smith	AR-OK	124	Hartford-	CT
105	Fort Walton Beach	FL		New Britain-	
106	Fort Wayne	IN		Middletown-	
107	Fort Worth-	TX		Bristol	
	Arlington		125	Hickory	NC
108	Fresno	CA	126	Honolulu	HI
109	Gadsden	AL	127	Houma-	LA
110	Gainesville	FL		Thibodaux	
111	Galveston-	TX	128	Houston	TX

Table 10.1 (Continued)

129	Huntington- Ashland	WV-KY-OH	153	Lake Charles	LA
			154	Lake County	IL
130	Huntsville	AL	155	Lakeland- Winter Haven	FL
131	Indianapolis	IN			
132	Iowa City	IA	156	Lancaster	PA
133	Jackson	MI	157	Lansing- East Lansing	MI
134	Jackson	MS			
135	Jacksonville	FL	158	Laredo	TX
136	Jacksonville	NC	159	Las Cruces	NM
137	Janesville- Beloit	WI	160	Las Vegas	NV
			161	Lawrence	KS
138	Jersey City	NJ	162	Lawton	OK
139	Johnson City- Kingsport- Bristol	TN-VA	163	Lewiston- Auburn	ME
			164	Lexington- Fayette	KY
140	Johnstown	PA			
141	Joliet	IL	165	Lima	OH
142	Joplin	MO	166	Lincoln	NE
143	Kalamazoo	MI	167	Little Rock- North Little Rock	AR
144	Kankakee	IL			
145	Kansas City	MO-KS	168	Longview- Marshall	TX
146	Kenosha	WI			
147	Killeen- Temple	TX	169	Lorain- Elyria	OH
148	Knoxville	TN	170	Los Angeles- Long Beach	CA
149	Kokomo	IN			
150	La Crosse	WI	171	Louisville	KY-IN
151	Lafayette	LA	172	Lubbock	TX
152	Lafayette- West Lafayette	IN	173	Lynchburg	VA
			174	Macon-	GA

Table 10.1 (Continued)

	Warner Robins		194	Naples	FL
175	Madison	WI	195	Nashville	TN
176	Manchester- Nashua	NH	196	Nassau- Suffolk	NY
177	Mansfield	OH	197	New Bedford- Fall River- Attleboro	MA
178	McAllen- Edinburg- Mission	TX	198	New Haven- Meriden- Waterburg	CT
179	Medford	OR			
180	Melbourne- Titusville- Palm Bay	FL	199	New London- Norwich	CT-RI
181	Memphis	TN-AR-MS	200	New Orleans	LA
182	Miami- Hialeah	FL	201	New York	NY
			202	Newark	NJ
183	Middlesex- Somerset- Hunterdon	NJ	203	Niagara Falls	NY
			204	Norfolk- Virginia Beach- Newport News	VA
184	Midland	TX			
185	Milwaukee	WI	205	Oakland	CA
186	Minneapolis- St. Paul	MN-WI	206	Ocala	FL
			207	Odessa	TX
187	Mobile	AL	208	Oklahoma City	OK
188	Modesto	CA	209	Olympia	WA
189	Monmouth- Ocean	NJ	210	Omaha	NE-IA
			211	Orange County	NY
190	Monroe	LA	212	Orlando	FL
191	Montgomery	AL	213	Owensboro	KY
192	Muncie	IN	214	Oxnard- Ventura	CA
193	Muskegon	MI			

Table 10.1 (Continued)

215	Panama City	FL		Kennewick-	
216	Parkersburg-	WV-OH		Pasco	
	Marietta		238	Richmond-	VA
217	Pascagoula	MS		Petersburg	
218	Pensacola	Fl	239	Riverside-	CA
219	Peoria	IL		San Bernardino	
220	Philadelphia	PA-NJ	240	Roanoke	VA
221	Phoenix	AZ	241	Rochester	MN
222	Pine Bluff	AR	242	Rochester	NY
223	Pittsburgh	PA	243	Rockford	IL
224	Pittsfield	MA	244	Sacramento	CA
225	Portland	ME	245	Saginaw-	MI
226	Portland	OR		Bay City-	
227	Portsmouth-	NH		Midland	
	Dover-		246	St. Cloud	MN
	Rochester		247	St. Joseph	MO
228	Poughkeepsie	NY	248	St. Louis	MO-IL
229	Providence-	RI	249	Salem	OR
	Pawtucket-		250	Salinas-	CA
	Woonsocket			Seaside-	
230	Provo-	UT		Monterey	
	Orem		251	Salt Lake City-	UT
231	Pueblo	CO		Ogden	
232	Racine	WI	252	San Angelo	TX
233	Raleigh-	NC	253	San Antonio	TX
	Durham		254	San Diego	CA
234	Reading	PA	255	San Francisco	CA
235	Redding	CA	256	San Jose	CA
236	Reno	NV	257	Santa Barbara-	CA
237	Richland-	WA		Santa Maria-	

Table 10.1 (Continued)

	Lompoc		281	Tallahassee	FL
258	Santa Cruz	CA	282	Tampa-	FL
259	Santa Fe	NM		St. Petersburg-	
260	Santa Rosa-	CA		Clearwater	
	Petaluma		283	Terre Haute	IN
261	Sarasota	FL	284	Texarkana	TX-AR
262	Savannah	GA	285	Toledo	OH
263	Scranton-	PA	286	Topeka	KS
	Wilkes-		287	Trenton	NJ
	Barre		288	Tucson	AZ
264	Seattle	WA	289	Tulsa	OK
265	Sharon	PA	290	Tuscaloosa	AL
266	Sheboygan	WI	291	Tyler	TX
267	Sherman-	TX	292	Utica-	NY
	Denison			Rome	
268	Shreveport	LA	293	Vallejo-	CA
269	Sioux City	IA-NE		Fairfield-	
270	Sioux Falls	SD		Nape	
271	South Bend-	IN	294	Vancouver	WA
	Mishawaka		295	Victoria	TX
272	Spokaane	WA	296	Vineland-	NJ
273	Springfield	IL		Millville-	
274	Springfield	MO		Bridgeton	
275	Springfield	MA	297	Visalia-	CA
276	State College	PA		Tulare-	
277	Steubenville-	OH-WV		Porterville	
	Weirton		298	Waco	TX
278	Stockton	CA	299	Washington	DC-MD-VA
279	Syracuse	NY	300	Waterloo-	IA
280	Tacoma	WA		Cedar Falls	

Table 10.1 (Continued)

301	Wausau	WI	331	NCR(Jefferson City)	MO
302	West Palm Beach- Boca Raton- Delray Beach	FL	332	NCR(Kahulai)	HI
			333	NCR(Kosciuske)	MS
			334	NCR(Leesburg)	FL
303	Wheeling	WV-OH	335	NCR(Lewisburg)	PA
304	Wichita	KS	336	NCR(Lewistown)	MT
305	Wichita Falls	TX	337	NCR(Marlinton)	WV
306	Williamsport	PA	338	NCR(Manshfielsd)	WI
307	Wilmington	DE-NJ-MD	339	NCR(Maced)	CA
308	Wilmington	NC	340	NCR(Milledgeville)	GA
309	Worcester	MA	341	NCR(Minot)	ND
310	Yakima	WA	342	NCR(Montpelier)	VT
311	York	PA	343	NCR(Mt. Vermon)	OH
312	Youngstown- Warren	OH	344	NCR(Newport)	RI
			345	NCR(Oneonta)	NY
313	Yuba City	CA	346	NCR(Opelouses)	LA
314	NCR(Ames)	IA	347	NCR(Pierre)	SD
315	NCR(Ansley)	NE	348	NCR(Plymouth)	NH
316	NCR(Appomattox)	VA	349	NCR(Richfield)	UT
317	NCR(Bend)	OR	350	NCR(Riverton)	WY
318	NCR(Brady)	TX	351	NCR(Salida)	CO
319	NCR(Brainerd)	MN	352	NCR(Salmon)	ID
320	NCR(Clare)	MI	353	NCR(Sanford)	NC
321	NCR(Columbia)	TN	354	NCR(Skowhegan)	ME
322	NCR(Danilk)	KY	355	NCR(Socorrs)	NM
323	NCR(Dover)	DE	356	NCR(Springerville)	AZ
324	NCR(Easton)	MD	357	NCR(Stillwater)	OK
325	NCR(Effingham)	IL	358	NCR(Sumter)	SC
326	NCR(Fairbanks)	AK	359	NCR(Sylacauga)	AL
327	NCR(Great Bend)	KS	360	NCR(Tonopah)	NV
328	NCR(Greencastle)	IN	361	NCR(Weuatchee)	WA
329	NCR(Greenfield)	MA	362	NCR(Willimatic)	CT
330	NCR(Hot Spring)	AR			

11. APPENDIX B: REAL INCOME DEFLATION FACTORS

Source : Consumer Price Index (CPI) Detailed Reports (Table 25A), Jan. 1980.

Deflator : CPI for all urban consumers, CPI(U), annual average by regions across city sizes.

City size: $A = Population > 1,250,000$.
$B = 385,000 \leq Population \leq 1,250,000$.
$C = 75,000 \leq Population \leq 385,000$.
$D = Population < 75,000$ or Rural areas.

Notes 1: Fixed price pattern across regions and city size over time is assumed in this study.

Notes 2: According to the CPI Detailed Report (Jan.,1989), CPI(U) reported for Dec. of the years 1959, 1969 and 1979 are as follows:

YEAR	CPI(U)[a]	CPI(U)[b]
1977	62.1	100.0
1969	37.7	60.7
1959	29.4	47.3

[a]Based on 1985=100.

[b]Based on 1977=100.

Table 11.1: CPI(U) Deflator by Year, Region and City Size

		Year			
Region	City size	1959	1969	1977	1979
.	A	55.30	70.91	100.0	116.8
SOUTH	B	55.82	71.58	100.0	117.9
.	C	55.86	71.64	100.0	118.0
.	D	55.01	70.54	100.0	116.2
.	A	55.20	70.79	100.0	116.6
WEST	B	56.24	72.12	100.0	118.8
.	C	55.63	71.33	100.0	117.5
.	D	55.30	70.91	100.0	116.8
.	A	56.24	72.12	100.0	118.8
NORTH CENTRAL	B	56.01	71.82	100.0	118.3
.	C	55.53	71.21	100.0	117.3
.	D	55.53	71.21	100.0	117.3
.	A	53.78	68.96	100.0	113.6
NORTH EAST	B	54.87	70.36	100.0	115.9
.	C	56.01	71.82	100.0	118.3
.	D	54.68	70.12	100.0	115.5

Table 11.2: Regional Price Index

	Region Size			
Region	A	B	C	D
SOUTH	1.6133	1.3308	1.3370	1.00
WEST	1.4686	1.2141	1.3370	1.00
NORTH CENTRAL	1.2576	1.0672	1.3370	1.00
NORTH EAST	1.2353	0.9902	1.3370	1.00

12. APPENDIX C: 12 CRs AND THEIR LINK-CITIES

CR	MSA	STATE	LINK-CITY	ADJUSTMENT
30	Beaver County	PA	Rochester	N/A
32	Banton Harbour	MI	St. Joseph	N/A
33	Berger-Passaic	NJ	Paterson	N/A
44	Bradenton	FL	Tampa	Adjustment 1
45	Brazoria	TX	Freeport	N/A
105	Fort Walton Beach	FL	Crestview	N/A
126	Honolulu	HI	San Francisco Los Angeles	Adjustment 2
154	Lake County	IL	Waukegan	N/A
179	Medford	OR	Grants Pass	N/A
211	Orange County	NY	Newburgh	N/A
265	Sharon	PA	Mercer	N/A
332	NCR(Kahulai)	HI	Honolulu	Adjustment 3

ADJUSTMENT 1: If CRs locate in the north (south) of Bradenton,

Mileage(Bradenton) = Mileage(Tampa) + (-) 33.

ADJUSTMENT 2: If CRs have shorter mileages to San Francisco than to Los Angeles, then,

Mileage(Honolulu)= Mileage(San Francisco) + 2397;

if CRs have shorter mileages to Los Angeles, then,

Mileage(Honolulu)= Mileage(Los Angeles) + 2563,

where 2397 and 2563 are the direct flight distances (in miles) from San Francisco and Los Angeles to Honolulu respectively.

ADJUSTMENT 3: Mileage(Kahulai) = Mileage(Honolulu) + 131.

N/A : No adjustment (the link-city is located inside the CR).

13. APPENDIX D: INTRA-CR-INTER-COUNTY DISTANCES

Table 13.1: Intra-CR-Inter-County Distances

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
2	1-2	18		2-3	19		2-13	14
3	1-2	19		2-4	23		2-14	17
4	1-2	28		3-4	22		2-15	68
	1-3	41	19	1-2	49		2-16	25
	1-4	27		1-3	56		2-17	19
	1-5	39		1-4	53		2-18	34
	1-6	17		1-5	49		3-4	51
	2-3	51		1-6	83		3-5	19
	2-4	55		1-7	38		3-6	61
	2-5	67		1-8	70		3-7	35
	2-6	45		1-9	65		3-8	39
	3-4	55		1-10	38		3-9	56
	3-5	36		1-11	56		3-10	18
	3-6	24		1-12	19		3-11	34
	4-5	36		1-13	45		3-12	37
	4-6	31		1-14	30		3-13	56
	5-6	22		1-15	64		3-14	58
7	1-2	21		1-16	29		3-15	32
	1-3	24		1-17	66		3-16	47
	1-4	38		1-18	15		3-17	70
	2-3	18		2-3	69		3-18	58
	2-4	37		2-4	28		4-5	31
	3-4	19		2-5	59		4-6	30
9	1-2	30		2-6	49		4-7	15
16	1-2	27		2-7	34		4-8	26
	1-3	21		2-8	54		4-9	12
	2-3	26		2-9	32		4-10	48
18	1-2	15		2-10	65		4-11	17
	1-3	14		2-11	45		4-12	34
	1-4	8		2-12	46		4-13	14

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	4-14	29		7-8	32		10-11	37
	4-15	40		7-9	27		10-12	19
	4-16	19		7-10	33		10-13	54
	4-17	19		7-11	18		10-14	47
	4-18	38		7-12	19		10-15	47
	5-6	40		7-13	21		10-16	40
	5-7	25		7-14	24		10-17	65
	5-8	20		7-15	43		10-18	38
	5-9	37		7-16	13		11-12	37
	5-10	32		7-17	32		11-13	31
	5-11	14		7-18	28		11-14	42
	5-12	30		8-9	25		11-15	25
	5-13	45		8-10	52		11-16	31
	5-14	49		8-11	14		11-17	36
	5-15	15		8-12	51		11-18	42
	5-16	38		8-13	40		12-13	37
	5-17	50		8-14	51		12-14	28
	5-18	49		8-15	20		12-15	45
	6-7	45		8-16	43		12-16	21
	6-8	23		8-17	41		12-17	52
	6-9	18		8-18	60		12-18	21
	6-10	69		9-10	59		13-14	19
	6-11	32		9-11	22		13-15	54
	6-12	64		9-12	46		13-16	16
	6-13	38		9-13	20		13-17	15
	6-14	57		9-14	39		13-18	36
	6-15	43		9-15	45		14-15	67
	6-16	39		9-16	31		14-16	11
	6-17	30		9-17	16		14-17	34
	6-18	58		9-18	50		14-18	17

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	15-16	56		3-6	61	38	1-2	38
	15-17	61		3-7	28	42	1-2	28
	15-18	65		4-5	39		1-3	36
	16-17	31		4-6	37		1-4	51
	16-18	19		4-7	23		1-5	20
	17-18	51		5-6	49		2-3	21
20	1-2	28		5-7	16		2-4	48
21	1-2	14		6-7	35		2-5	17
	1-3	17	27	1-2	23		3-4	27
	1-4	34		1-3	20		3-5	16
	2-3	26		1-4	29		4-5	31
	2-4	48		2-3	22	52	1-2	27
	3-4	25		2-4	14	53	1-2	24
22	1-2	25		3-4	36	59	1-2	36
23	1-2	24	29	1-2	32		1-3	27
	1-3	52		1-3	33		2-3	28
	2-3	28		2-3	23	58	1-2	30
25	1-2	32	33	1-2	16	59	1-2	34
	1-3	42	35	1-2	24		1-3	36
	1-4	44	36	1-2	23		1-4	18
	1-5	21	37	1-2	37		1-5	17
	1-6	30		1-3	23		1-6	25
	1-7	21		1-4	51		1-7	47
	2-3	19		1-5	44		2-3	14
	2-4	20		2-3	34		2-4	22
	2-5	19		2-4	24		2-5	42
	2-6	46		2-5	31		2-6	40
	2-7	11		3-4	37		2-7	20
	3-4	39		3-5	55		3-4	29
	3-5	21		4-5	55		3-5	38

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	3-6	49		1-3	48	69	1-2	27
	3-7	34		2-3	39	70	1-2	12
	4-5	29	64	1-2	24		1-3	24
	4-6	20		1-3	27		2-3	26
	4-7	26		1-4	30	71	1-2	68
	5-6	42		1-5	13		1-3	38
	5-7	56		1-6	20		1-4	41
	6-7	36		1-7	43		1-5	29
60	1-2	21		2-3	25		1-6	67
	1-3	20		2-4	18		1-7	17
	1-4	7		2-5	18		2-3	30
	2-3	31		2-6	17		2-4	29
	2-4	15		2-7	22		2-5	46
	3-4	18		3-4	43		2-6	22
61	1-2	25		3-5	33		2-7	62
	1-3	19		3-6	38		3-4	34
	1-4	18		3-7	47		3-5	23
	1-5	28		4-5	18		3-6	25
	1-6	30		4-6	11		3-7	32
	2-3	20		4-7	18		4-5	57
	2-4	32		5-6	7		4-6	51
	2-5	19		5-7	33		4-7	55
	2-6	31		6-7	45		5-6	35
	3-4	34	65	1-2	21		5-7	26
	3-5	33	66	1-2	23		6-7	57
	3-6	42		1-3	25	72	1-2	22
	4-5	22		1-4	25	73	1-2	20
	4-6	17		2-3	16	74	1-2	34
	5-6	14		2-4	41		1-3	34
62	1-2	19		3-4	50		1-4	62

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	1-5	44		3-5	15		1-3	30
	1-6	20		4-5	24		1-4	24
	2-3	36	81	1-2	24		2-3	16
	2-4	28		1-3	44		2-4	17
	2-5	32		2-3	20		3-4	24
	2-6	24	82	1-2	49	94	1-2	38
	3-4	64		1-3	32	98	1-2	15
	3-5	68		1-4	84	103	1-2	22
	3-6	54		1-5	32	104	1-2	27
	4-5	36		1-6	30		1-3	28
	4-6	52		1-7	56		2-3	34
	5-6	24		2-3	51	106	1-2	27
75	1-2	17		2-4	50		1-3	23
76	1-2	33		2-5	27		2-3	37
	1-3	12		2-6	73	107	1-2	36
	2-3	21		2-7	38		1-3	36
77	1-2	18		3-4	60		2-3	38
	1-3	27		3-5	24	110	1-2	24
	1-4	26		3-6	28	112	1-2	16
	2-3	33		3-7	32	113	1-2	35
	2-4	18		4-5	46	115	1-2	26
	3-4	20		4-6	88	119	1-2	19
80	1-2	15		4-7	28		1-3	20
	1-3	34		5-6	46		1-4	29
	1-4	47		5-7	24		1-5	24
	1-5	49		6-7	60		1-6	38
	2-3	34	83	1-2	22		1-7	37
	2-4	38	85	1-2	81		2-3	20
	2-5	48	86	1-2	23		2-4	41
	3-4	27	93	1-2	14		2-5	43

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	2-6	36		1-4	68		2-3	23
	2-7	18		1-5	35		2-4	31
	3-4	25		2-3	41		2-5	42
	3-5	39		2-4	33		2-6	21
	3-6	18		2-5	39		2-7	47
	3-7	23		3-4	43		2-8	40
	4-5	26		3-5	80		3-4	38
	4-6	32		4-5	38		3-5	28
	4-7	48	129	1-2	20		3-6	19
	5-6	57		1-3	24		3-7	42
	5-7	61		1-4	43		3-8	19
	6-7	27		1-5	38		4-5	29
120	1-2	19		1-6	20		4-6	19
	1-3	27		2-3	20		4-7	20
	2-3	44		2-4	34		4-8	40
123	1-2	27		2-5	37		5-6	21
	1-3	44		2-6	29		5-7	19
	1-4	16		3-4	19		5-8	17
	2-3	20		3-5	17		6-7	26
	2-4	23		3-6	16		6-8	26
	3-4	21		4-5	14		7-8	37
124	1-2	21		4-6	31	134	1-2	31
	1-3	15		5-6	22		1-3	29
	-3	23	131	1-2	22		2-3	24
125	1-2	32		1-3	40	135	1-2	30
	1-3	18		1-4	20		1-3	47
	2-3	29		1-5	42		1-4	28
127	1-2	33		1-6	25		2-3	22
128	1-2	35		1-7	40		2-4	36
	1-3	72		1-8	51		3-4	58

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
139	1-2	41	142	1-2	21		4-10	53
	1-3	17	145	1-2	43		5-6	41
	1-4	18		1-3	22		5-7	34
	1-5	18		1-4	42		5-8	17
	1-6	37		1-5	53		5-9	53
	1-7	32		1-6	53		5-10	22
	1-8	22		1-7	29		6-7	55
	2-3	31		1-8	53		6-8	56
	2-4	35		1-9	27		6-9	70
	2-5	25		1-10	39		6-10	45
	2-6	32		2-3	20		7-8	32
	2-7	55		2-4	36		7-9	20
	2-8	43		2-5	20		7-10	18
	3-4	25		2-6	23		8-9	46
	3-5	18		2-7	36		8-10	20
	3-6	20		2-8	29		9-10	42
	3-7	26		2-9	54	147	1-2	32
	3-8	14		2-10	23	148	1-2	34
	4-5	11		3-4	32		1-3	39
	4-6	41		3-5	34		1-4	43
	4-7	46		3-6	33		1-5	16
	4-8	36		3-7	25		1-6	44
	5-6	30		3-8	39		1-7	23
	5-7	42		3-9	38		2-3	47
	5-8	30		3-10	24		2-4	37
	6-7	33		4-5	56		2-5	22
	6-8	23		4-6	23		2-6	24
	7-8	12		4-7	57		2-7	44
140	1-2	37		4-8	67		3-4	16
141	1-2	25		4-9	65		3-5	31

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
172	1-2	24		1-11	40		5-11	56
	1-3	17		2-3	63		6-7	38
	1-4	19		2-4	35		6-8	63
	2-3	41		2-5	21		6-9	41
	2-4	13		2-6	57		6-10	42
	3-4	36		2-7	37		6-11	54
181	1-2	26		2-8	18		7-8	31
	1-3	24		2-9	63		7-9	12
	1-4	24		2-10	25		7-10	43
	2-3	39		2-11	67		7-11	32
	2-4	50		3-4	55		8-9	41
	3-4	29		3-5	43		8-10	41
183	1-2	30		3-6	21		8-11	58
	1-3	15		3-7	34		9-10	52
	2-3	15		3-8	63		9-11	22
185	1-2	26		3-9	31		10-11	74
	1-3	28		3-10	55	187	1-2	31
	1-4	20		3-11	37	189	1-2	28
	2-3	15		4-5	29	191	1-2	28
	2-4	29		4-6	63		1-3	31
	3-4	23		4-7	23		2-3	26
186	1-2	42		4-8	21	195	1-2	16
	1-3	22		4-9	26		1-3	17
	1-4	41		4-10	53		1-4	22
	1-5	23		4-11	41		1-5	49
	1-6	22		5-6	41		1-6	37
	1-7	18		5-7	18		1-7	31
	1-8	43		5-8	23		1-8	45
	1-9	23		5-9	28		2-3	28
	1-10	38		5-10	26		2-4	25

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	2-5	34		3-5	53		5-8	33
	2-6	31		3-6	51		6-7	52
	2-7	19		4-5	20		6-8	53
	2-8	31		4-6	60		7-8	13
	3-4	39		5-6	45	202	1-2	14
	3-5	58	201	1-2	20		1-3	35
	3-6	54		1-3	5		1-4	11
	3-7	33		1-4	38		2-3	21
	3-8	59		1-5	10		2-4	17
	4-5	53		1-6	30		3-4	38
	4-6	22		1-7	22	204	1-2	15
	4-7	43		1-8	23		1-3	16
	4-8	40		2-3	15		1-4	27
	5-6	45		2-4	58		1-5	30
	5-7	28		2-5	10		1-6	38
	5-8	25		2-6	10		1-7	38
	6-7	47		2-7	42		1-8	46
	6-8	25		2-8	43		1-9	62
	7-8	38		3-4	43		1-10	9
196	1-2	35		3-5	5		2-3	21
200	1-2	22		3-6	25		2-4	34
	1-3	29		3-7	27		2-5	35
	1-4	16		3-8	28		2-6	51
	1-5	36		4-5	48		2-7	45
	1-6	53		4-6	68		2-8	62
	2-3	20		4-7	22		2-9	69
	2-4	29		4-8	15		2-10	7
	2-5	39		5-6	20		3-4	13
	2-6	31		5-7	20		3-5	14
	3-4	42		5-7	32		3-6	30

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	3-7	24		2-3	47		1-8	43
	3-8	41		2-4	16		2-3	18
	3-9	48		2-5	22		2-4	27
	3-10	17		2-6	25		2-5	28
	4-5	8		3-4	60		2-6	55
	4-6	24		3-5	25		2-7	42
	4-7	18		3-6	56		2-8	36
	4-8	29		4-5	35		3-4	19
	4-9	42		4-6	31		3-5	13
	4-10	27		5-6	35		3-6	38
	5-6	121	210	1-2	12		3-7	24
	5-7	10		1-3	18		3-8	20
	5-8	21		1-4	29		4-5	16
	5-9	34		2-3	30		4-6	42
	5-10	30		2-4	33		4-7	34
	6-7	6		3-4	35		4-8	35
	6-8	22	212	1-2	34		5-6	28
	6-9	19		1-3	16		5-7	18
	6-10	42		2-3	45		5-8	21
	7-8	17	216	1-2	20		6-7	17
	7-9	24	218	1-2	27		6-8	27
	7-10	40	219	1-2	23		7-8	10
	8-9	34		1-3	29	223	1-2	39
	8-10	57		2-3	23		1-3	24
	9-10	64	220	1-2	39		1-4	27
205	1-2	37		1-3	32		2-3	36
	1-3	38		1-4	17		2-4	28
	1-4	44		1-5	22		3-4	41
	1-5	30		1-6	37	226	1-2	24
	1-6	62		1-7	38		1-3	45

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	1-4	53		2-3	22		5-6	22
	2-3	28		2-4	27		5-7	30
	2-4	46		2-5	28		5-8	27
	3-4	26		2-6	21		5-9	50
227	1-2	17		2-7	31		5-10	36
229	1-2	12		2-8	25		5-11	35
	1-3	16		2-9	21		5-12	42
	1-4	21		2-10	12		5-13	17
	2-3	10		2-11	10		6-7	16
	2-4	13		2-12	14		6-8	34
	3-4	22		2-13	13		6-9	23
233	1-2	35		3-4	46		6-10	17
	1-3	12		3-5	50		6-11	14
	1-4	20		3-6	34		6-12	23
	2-3	47		3-7	47		6-13	8
	2-4	29		3-8	39		7-8	50
	3-4	32		3-9	23		7-9	24
237	1-2	35		3-10	19		7-10	28
238	1-2	30		3-11	25		7-11	22
	1-3	38		3-12	16		7-12	31
	1-4	52		3-13	35		7-13	23
	1-5	65		4-5	21		8-9	51
	1-6	19		4-6	33		8-10	35
	1-7	11		4-7	46		8-11	41
	1-8	50		4-8	10		8-12	41
	1-9	16		4-9	54		8-13	27
	1-10	21		4-10	38		9-10	11
	1-11	16		4-11	41		9-11	9
	1-12	25		4-12	44		9-12	10
	1-13	27		4-13	25		9-13	28

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	10-11	6	246	1-2	20		4-8	33
	10-12	6		1-3	30		4-9	29
	10-13	19		2-3	38		4-10	27
	11-12	9	248	1-2	31		5-6	44
	11-13	19		1-3	31		5-7	30
	12-13	25		1-4	38		5-8	23
239	1-2	87		1-5	47		5-9	24
240	1-2	24		1-6	90		5-10	31
	1-3	19		1-7	61		6-7	60
	2-3	9		1-8	71		6-8	35
242	1-2	31		1-9	47		6-9	46
	1-3	27		1-10	59		6-10	31
	1-4	44		2-3	38		7-8	25
	1-5	46		2-4	29		7-9	53
	2-3	25		2-5	38		7-10	43
	2-4	34		2-6	65		8-9	40
	2-5	27		2-7	57		8-10	23
	3-4	59		2-8	53		9-10	18
	3-5	21		2-9	19	249	1-2	34
	4-5	61		2-10	34	251	1-2	27
243	1-2	28		3-4	17		1-3	22
244	1-2	27		3-5	29		2-3	42
	1-3	48		3-6	72	253	1-2	30
	1-4	74		3-7	30		1-3	37
	2-3	55		3-8	46		2-3	19
	2-4	70		3-9	43	255	1-2	26
	3-4	35		3-10	43		1-3	45
245	1-2	20		4-5	12		2-3	19
	1-3	26		4-6	57	259	1-2	28
	2-3	29		4-7	28	262	1-2	28

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
263	1-2	44	284	1-2	33		2-5	53
	1-3	23	285	1-2	29		2-6	34
	1-4	53		1-3	29		2-7	37
	1-5	36		2-3	20		2-8	41
	2-3	24	289	1-2	49		2-9	70
	2-4	31		1-3	56		2-10	50
	2-5	20		1-4	34		2-11	49
	3-4	34		1-5	46		2-12	34
	3-5	24		2-3	43		2-13	41
	4-5	48		2-4	34		3-4	68
264	1-2	40		2-5	77		3-5	42
268	1-2	20		3-4	22		3-6	22
269	1-2	25		3-5	26		3-7	24
273	1-2	19		4-5	43		3-8	24
274	1-2	25	292	1-2	26		3-9	53
275	1-2	16	293	1-2	27		3-10	29
277	1-2	11	299	1-2	34		3-11	28
	1-3	13		1-3	25		3-12	20
	2-3	16		1-4	44		3-13	27
279	1-2	24		1-5	19		4-5	27
	1-3	41		1-6	10		4-6	53
	2-3	27		1-7	7		4-7	45
281	1-2	24		1-8	15		4-8	45
282	1-2	47		1-9	38		4-9	29
	1-3	20		1-10	30		4-10	55
	1-4	54		1-11	42		4-11	72
	2-3	27		1-12	25		4-12	49
	2-4	33		1-13	11		4-13	44
	3-4	37		2-3	22		5-6	28
283	1-2	15		2-4	78		5-7	18

Table 13.1 (Continued)

CR	County	D_{ij}	CR	County	D_{ij}	CR	County	D_{ij}
	5-8	21		10-13	23			
	5-9	24		11-12	35			
	5-10	35		11-13	36			
	5-11	50		12-13	8			
	5-12	22	300	1-2	16			
	5-13	17	303	1-2	18			
	6-7	15		1-3	20			
	6-8	22		2-3	20			
	6-9	48	304	1-2	26			
	6-10	37	307	1-2	17			
	6-11	43		1-3	14			
	6-12	13		2-3	31			
	6-13	19	311	1-2	26			
	7-8	9	312	1-2	20			
	7-9	34	313	1-2	20			
	7-10	25						
	7-11	37						
	7-12	4						
	7-13	5						
	8-9	30						
	8-10	17						
	8-11	30						
	8-12	9						
	8-13	7						
	9-10	32						
	9-11	49						
	9-12	37						
	9-13	30						
	10-11	17						
	10-12	24						

14. APPENDIX E: OWN DISTANCES

Table 14.1: Own Distances

CR	D_{ii}	CR	D_{ii}	CR	D_{ii}
1	22	31	61	61	35
2	22	32	23	62	30
3	23	33	17	63	33
4	40	34	41	64	37
5	28	35	29	65	31
6	27	36	31	66	30
7	29	37	44	67	34
8	20	38	52	68	21
9	32	39	15	69	28
10	22	40	24	70	26
11	37	41	27	71	42
12	17	42	28	72	28
13	23	43	20	73	28
14	19	44	25	74	54
15	19	45	27	75	24
16	29	46	19	76	34
17	20	47	16	77	30
18	25	48	26	78	35
19	49	49	21	79	17
20	23	50	23	80	59
21	42	51	19	81	34
22	26	52	19	82	60
23	39	53	25	83	32
24	83	54	58	84	19
25	44	55	19	85	93
26	15	56	23	86	29
27	32	57	44	87	27
28	19	58	29	88	15
29	41	59	41	89	15
30	16	60	27	90	23

Table 14.1 (Continued)

CR	D_{ii}	CR	D_{ii}	CR	D_{ii}
91	24	121	23	151	28
92	61	122	16	152	16
93	31	123	39	153	28
94	41	124	26	154	16
95	20	125	30	155	38
96	24	126	23	156	20
97	19	127	38	157	30
98	29	128	54	158	51
99	24	129	35	159	50
100	40	130	20	160	71
101	28	131	39	161	16
102	20	132	19	162	24
103	30	133	19	163	12
104	36	134	40	164	28
105	23	135	47	165	25
106	29	136	19	166	22
107	38	137	19	167	42
108	70	138	8	168	30
109	17	139	47	169	18
110	32	140	42	170	117
111	24	141	31	171	36
112	26	142	26	172	22
113	36	143	17	173	28
114	30	144	22	174	33
115	29	145	52	175	25
116	41	146	14	176	23
117	53	147	35	177	18
118	18	148	37	178	29
119	43	149	18	179	43
120	42	150	15	180	37

Table 14.1 (Continued)

CR	D_{ii}	CR	D_{ii}	CR	D_{ii}
181	37	211	21	241	19
182	37	212	48	242	45
183	26	213	16	243	22
184	22	214	33	244	67
185	35	215	25	245	31
186	53	216	25	246	43
187	40	217	20	247	17
188	35	218	33	248	63
189	32	219	28	249	50
190	17	220	40	250	57
191	36	221	75	251	37
192	14	222	26	252	30
193	19	223	43	253	38
194	37	224	26	254	53
195	44	225	18	255	46
196	52	226	53	256	28
197	20	227	23	257	39
198	15	228	23	258	22
199	17	229	20	259	14
200	53	230	42	260	39
201	40	231	36	261	23
202	28	232	15	262	28
203	18	233	34	263	37
204	43	234	24	264	49
205	30	235	54	265	19
206	30	236	97	266	15
207	22	237	50	267	27
208	48	238	39	268	33
209	24	239	120	269	30
210	38	240	27	270	18

Table 14.1 (Continued)

CR	D_{ii}	CR	D_{ii}	CR	D_{ii}
271	17	302	35	333	174
272	33	303	25	334	308
273	24	304	39	335	159
274	29	305	19	336	314
275	19	306	30	337	139
276	33	307	32	338	180
277	19	308	17	339	409
278	28	309	29	340	192
279	40	310	51	341	207
280	39	311	33	342	91
281	28	312	25	343	148
282	42	313	33	344	28
283	20	314	170	345	200
284	36	315	250	346	177
285	35	316	238	347	202
286	19	317	230	348	95
287	12	318	403	349	217
288	90	319	237	350	225
289	58	320	240	351	238
290	24	321	244	352	280
291	26	322	210	353	240
292	44	323	53	354	160
293	30	324	128	355	161
294	19	325	195	356	245
295	22	326	166	357	261
296	16	327	224	358	142
297	48	328	159	359	180
298	23	329	91	360	276
299	57	330	165	361	200
300	25	331	229	362	60
301	31	332	127		

15. APPENDIX F: VARIABLES AND POTENTIALS PLOTS

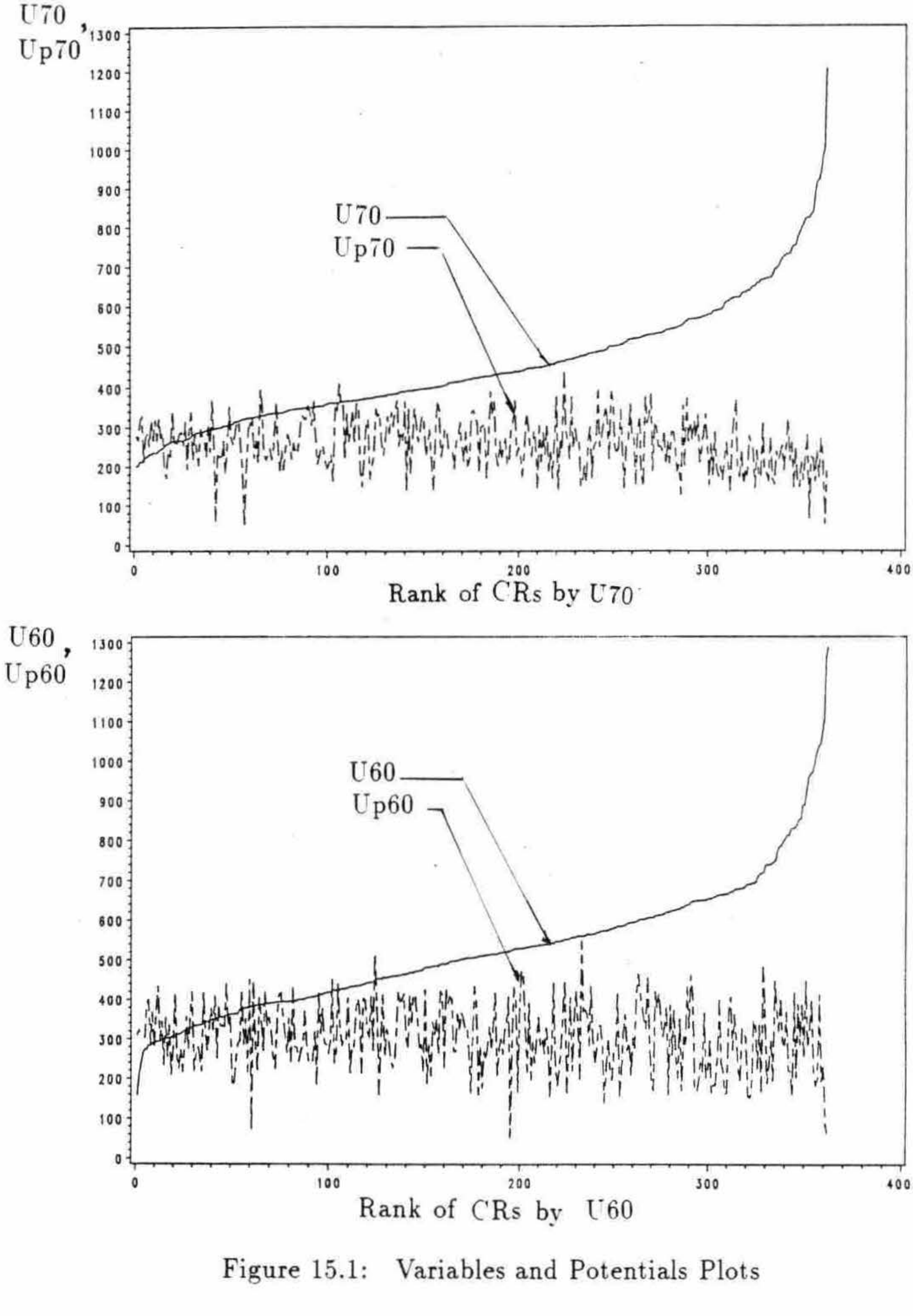
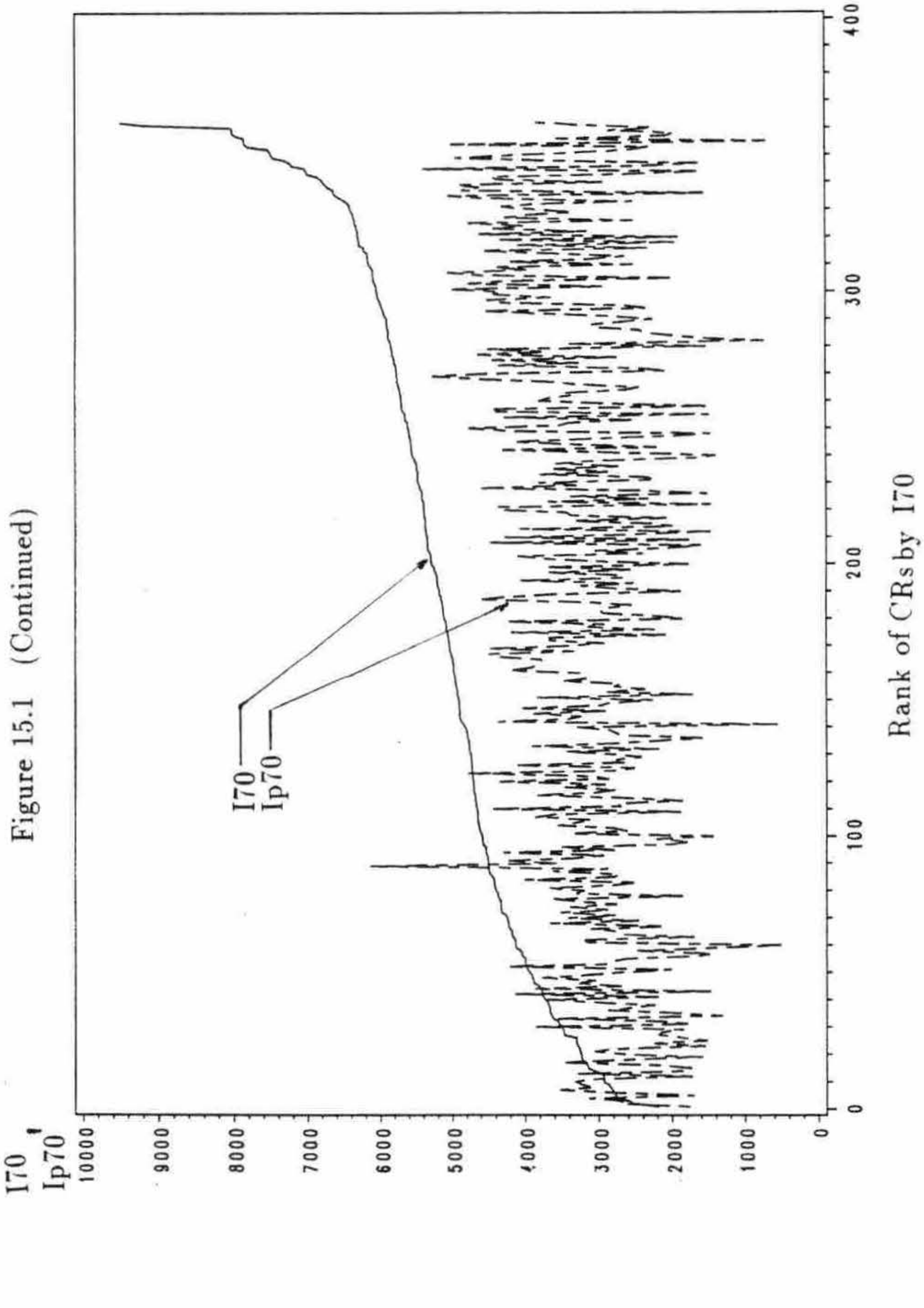


Figure 15.1: Variables and Potentials Plots

Figure 15.1 (Continued)



16. APPENDIX G: $\widehat{M80}$, $\widehat{M90}$ BY CRs

Table 16.1: $\widehat{M80}$ and $\widehat{M90}$ by CRs

CR	$\widehat{M80}$	$\widehat{M90}$	CR	$\widehat{M80}$	$\widehat{M90}$
1	28414	46245	31	29563	51587
2	83650	91251	32	23156	26713
3	26404	33374	33	157208	132516
4	87452	101190	34	26027	48216
5	89015	173394	35	47512	57685
6	22953	37854	36	32345	36414
7	58927	80429	37	84190	142502
8	11535	15610	38	10905	33481
9	34289	68115	39	34506	37197
10	609833	624083	40	26850	36878
11	104775	159731	41	37568	102956
12	16968	21165	42	384564	416764
13	13782	28036	43	85758	117606
14	72110	84807	44	38363	75684
15	20731	32545	45	39203	95403
16	40817	54646	46	35752	79570
17	21860	33915	47	105020	111966
18	24982	47212	48	17671	62109
19	337507	493118	49	23556	73023
20	33551	53899	50	88064	85524
21	59293	86134	51	14509	15849
22	48710	68733	52	27153	32295
23	105699	278170	53	42338	51610
24	76000	124914	54	15188	42565
25	224323	258642	55	33592	40017
26	24196	34816	56	54187	59059
27	59636	142492	57	84214	138675
28	18597	22111	58	26097	41922
29	49711	80480	59	114855	168936
30	16171	21865	60	17612	41103

Table 16.1 (Continued)

CR	$\widehat{M80}$	$\widehat{M90}$	CR	$\widehat{M80}$	$\widehat{M90}$
61	47165	80690	91	27641	35716
62	513631	595716	92	80450	133443
63	37072	80392	93	27100	43261
64	139499	172968	94	38256	52698
65	27370	55427	95	86326	88130
66	203128	172815	96	30621	44338
67	129269	174111	97	52690	54666
68	37030	45360	98	14154	24893
69	85969	126306	99	13309	22541
70	59027	65515	100	56103	113768
71	157993	191512	101	366086	518083
72	53977	105406	102	72000	153425
73	11349	13652	103	28705	101678
74	332367	541668	104	29315	47757
75	11502	14770	105	49114	50974
76	55218	69159	106	47388	49922
77	140090	133218	107	216203	298105
78	62851	132525	108	77370	141749
79	19340	23286	109	11665	17647
80	358803	615888	110	43160	84975
81	52351	73857	111	43526	64372
82	472364	465464	112	62372	67859
83	39685	35874	113	14255	19707
84	12605	15318	114	26802	32509
85	40313	53503	115	76232	100270
86	19525	36454	116	29128	30352
87	78885	145564	117	32585	65316
88	21728	25426	118	28385	37016
89	12484	14805	119	92818	138058
90	13844	21156	120	71270	122188

Table 16.1 (Continued)

CR	$\widehat{M80}$	$\widehat{M90}$	CR	$\widehat{M80}$	$\widehat{M90}$
121	14373	18750	151	25616	58144
122	38058	66174	152	31738	41017
123	57900	80126	153	23756	45362
124	123737	123632	154	98450	123067
125	21972	32032	155	55270	124871
126	261670	308120	156	35428	52630
127	19139	36276	157	81149	96798
128	428146	990028	158	11874	18545
129	31169	48502	159	19961	47009
130	58927	45639	160	144659	293837
131	147804	167957	161	24634	32054
132	29022	36326	162	42963	46909
133	21081	27119	163	13377	20957
134	38507	75238	164	39631	82041
135	138648	199204	165	15191	19656
136	51980	55649	166	44457	63982
137	20894	23661	167	67333	121411
138	33926	43500	168	23816	56127
139	37940	75419	169	36309	44413
140	19899	24387	170	1127713	1196996
141	47194	88565	171	90077	112061
142	22347	34386	172	54214	90207
143	39857	50020	173	16871	28060
144	13501	17520	174	43894	49878
145	198524	245414	175	78132	88627
146	15959	20819	176	41311	70549
147	71301	122287	177	16827	18327
148	51781	94096	178	23273	81184
149	13171	13944	179	35905	74482
150	16100	27375	180	139804	114299

Table 16.1 (Continued)

CR	$\widehat{M80}$	$\widehat{M90}$	CR	$\widehat{M80}$	$\widehat{M90}$
181	135456	178991	211	39586	54793
182	359383	409588	212	163277	338265
183	143833	134449	213	12381	13849
184	17032	43124	214	167180	226462
185	147129	167543	215	24248	38186
186	269020	324247	216	20812	28013
187	50641	100762	217	33338	41491
188	46984	91104	218	69363	98592
189	159734	182536	219	47517	59928
190	23109	36802	220	435720	395060
191	39377	72135	221	359558	811780
192	25410	25279	222	14310	20236
193	17088	21651	223	174057	206618
194	33525	70081	224	16767	18317
195	97799	183348	225	31508	49414
196	371358	290798	226	213884	350332
197	41360	56922	227	50475	87871
198	78888	94536	228	40008	46943
199	44725	49200	229	82652	88999
200	151021	248195	230	46169	117541
201	471311	448685	231	18263	31758
202	201925	188220	232	23540	25372
203	20858	27840	233	94311	146418
204	230065	264091	234	26662	41369
205	307024	402579	235	29928	67670
206	22701	75023	236	49694	131754
207	22023	51616	237	26424	99224
208	139679	223126	238	89315	120689
209	33298	79769	239	366318	657996
210	100426	124104	240	26152	36230

Table 16.1 (Continued)

CR	$\widehat{M80}$	$\widehat{M90}$	CR	$\widehat{M80}$	$\widehat{M90}$
241	24989	26708	271	30021	39643
242	112946	128912	272	77234	137894
243	42681	41984	273	23135	37960
244	223557	397072	274	37030	69340
245	40921	49700	275	66669	72020
246	22181	50194	276	25655	39689
247	11994	16914	277	12714	17196
248	243445	267102	278	58669	99672
249	60409	128613	279	84558	97132
250	86479	110018	280	166954	192064
251	121928	267988	281	40398	71512
252	19346	36364	282	387452	715762
253	193134	309032	283	20766	22377
254	484599	792191	284	21698	29953
255	267741	327847	285	63779	81894
256	362957	398407	286	35516	36476
257	105645	108083	287	37240	44272
258	46414	95716	288	133596	286239
259	14902	39788	289	95529	190139
260	70286	142878	290	25813	42304
261	61957	123069	291	21804	49656
262	33881	50311	292	38971	38078
263	47643	72445	293	76566	143773
264	457689	485241	294	36632	86480
265	13391	17763	295	12500	25763
266	12741	17936	296	11635	16201
267	23548	30594	297	38319	76240
268	51153	80850	298	28630	55530
269	19498	24096	299	526155	530326
270	21151	37157	300	21436	31189

Table 16.1 (Continued)

CR	$\widehat{M80}$	$\widehat{M90}$	CR	$\widehat{M80}$	$\widehat{M90}$
301	14353	25848	332	21190	79733
302	150644	313451	333	145860	262797
303	16604	20496	334	148108	347905
304	62143	105926	335	148283	221041
305	33454	40438	336	82944	163525
306	12246	15693	337	96211	184375
307	63560	71733	338	160394	250555
308	22265	33176	339	181294	492863
309	65009	84491	340	187092	310462
310	30984	56481	341	47431	76916
311	34728	57984	342	51779	75746
312	50341	49526	343	233264	285045
313	29970	40694	344	28566	17796
314	140738	198848	345	191565	256180
315	84145	127666	346	137696	223030
316	168722	308329	347	68517	106448
317	164887	344757	348	51089	92878
318	264845	547814	349	41026	112478
319	146811	211463	350	61320	205598
320	184474	267672	351	75461	206766
321	144050	267096	352	142922	267928
322	168201	301504	353	220830	386730
323	27650	39812	354	70956	130009
324	36397	59364	355	99156	215509
325	187251	254475	356	122930	291591
326	95174	139616	357	164370	294051
327	153276	228644	358	120459	211015
328	174397	224585	359	135158	212523
329	33364	58889	360	32668	80785
330	159059	303812	361	158365	296279
331	198614	300873	362	35431	44874